



SOFTWARE: Patent.pm  
 SEQ ID NO 407  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo Sapiens  
 FEATURE:  
 NAME/KEY: SIGNAL  
 LOCATION: -18...-1  
 US-09-978-360A-407

Query Match 100.0%; Score 1626; DB 11; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-160;  
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAYIILCCWAGLPCLATCLDHHPTGSRPTVPGPLHFSGY 60  
 DB 1 MKAPGRVLIIICSVFSAYIILCCWAGLPCLATCLDHHPTGSRPTVPGPLHFSGY 60  
 QY 61 SVDPGKPLVREPCRSQAVVSSSGMIGSGIGAEIDSAECVFRNNOAFTVGFEDVQGRST 120  
 DB 61 SVDPGKPLVREPCRSQAVVSSSGMIGSGIGAEIDSAECVFRNNOAFTVGFEDVQGRST 120  
 QY 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMYWGGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMYWGGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 QY 181 VYTFTEEMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVYVGVMSDGYCR 240  
 DB 181 VYTFTEEMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVYVGVMSDGYCR 240  
 QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAPRSARHFTTEKAVFSRMAKKRPVFAHPSWR 300  
 DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAPRSARHFTTEKAVFSRMAKKRPVFAHPSWR 300  
 QY 301 TE 302  
 DB 301 TE 302

## RESULT 2

US-10-315-664-75  
 Sequence 75; Application US/10315664  
 Publication No. US20030203377A1  
 GENERAL INFORMATION:  
 APPLICANT: Dumas Milne Edwards, J.B.  
 APPLICANT: Bougueleret, L.  
 TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal  
 TITLE OF INVENTION: Peptides  
 FILE REFERENCE: GENSET.050CP3  
 CURRENT APPLICATION NUMBER: US/10/315,664  
 CURRENT FILING DATE: 2002-12-09  
 PRIOR APPLICATION NUMBER: US/09/599,360  
 PRIOR FILING DATE: 2000-06-21  
 PRIOR APPLICATION NUMBER: 60/113,686  
 PRIOR FILING DATE: 1998-12-22  
 PRIOR APPLICATION NUMBER: 60/141,032  
 PRIOR FILING DATE: 1999-06-25  
 PRIOR APPLICATION NUMBER: 09/469,099  
 PRIOR FILING DATE: 1999-12-21  
 NUMBER OF SEQ ID NOS: 123  
 SOFTWARE: Patent.pm  
 SEQ ID NO 75  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo Sapiens  
 FEATURE:  
 NAME/KEY: SIGNAL  
 LOCATION: -18...-1  
 US-10-315-664-75

Query Match 100.0%; Score 1626; DB 14; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-160;

Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAYIILCCWAGLPCLATCLDHHPTGSRPTVPGPLHFSGY 60  
 DB 1 MKAPGRVLIIICSVFSAYIILCCWAGLPCLATCLDHHPTGSRPTVPGPLHFSGY 60  
 QY 61 SVDPGKPLVREPCRSQAVVSSSGMIGSGIGAEIDSAECVFRNNOAFTVGFEDVQGRST 120  
 DB 61 SVDPGKPLVREPCRSQAVVSSSGMIGSGIGAEIDSAECVFRNNOAFTVGFEDVQGRST 120  
 QY 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMYWGGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMYWGGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 QY 181 VYTFTEEMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVYVGVMSDGYCR 240  
 DB 181 VYTFTEEMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVYVGVMSDGYCR 240  
 QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAPRSARHFTTEKAVFSRMAKKRPVFAHPSWR 300  
 DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAPRSARHFTTEKAVFSRMAKKRPVFAHPSWR 300  
 QY 301 TE 302  
 DB 301 TE 302

## RESULT 3

US-09-895-298-66  
 Sequence 66; Application US/09895298  
 Publication No. US20030078405A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: 47 Human Secreted Proteins  
 FILE REFERENCE: P2035P1  
 CURRENT APPLICATION NUMBER: US/09/895,298  
 CURRENT FILING DATE: 2001-07-02  
 PRIOR APPLICATION NUMBER: 09/591,16  
 PRIOR FILING DATE: 2000-06-09  
 PRIOR APPLICATION NUMBER: PCT/US99/23950  
 PRIOR FILING DATE: 1999-12-16  
 PRIOR APPLICATION NUMBER: 60/113,006  
 PRIOR FILING DATE: 1998-12-18  
 PRIOR APPLICATION NUMBER: 60/112,809  
 PRIOR FILING DATE: 1998-12-17  
 NUMBER OF SEQ ID NOS: 231  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 66  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (237)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-895-298-66

Query Match 99.8%; Score 1622; DB 10; Length 302;  
 Best Local Similarity 99.7%; Pred. No. 6,2e-160;  
 Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAYIILCCWAGLPCLATCLDHHPTGSRPTVPGPLHFSGY 60  
 DB 1 MKAPGRVLIIICSVFSAYIILCCWAGLPCLATCLDHHPTGSRPTVPGPLHFSGY 60  
 QY 61 SVDPGKPLVREPCRSQAVVSSSGMIGSGIGAEIDSAECVFRNNOAFTVGFEDVQGRST 120  
 DB 61 SVDPGKPLVREPCRSQAVVSSSGMIGSGIGAEIDSAECVFRNNOAFTVGFEDVQGRST 120  
 QY 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMYWGGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMYWGGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180

QY 181 VYTFETEMMAVYCDQIFODETGKRNROSGSFLLSTGMFTMIALBELCEIIVVYGWSDSYCR 240  
| 181 VYTFETEMMAVYCDQIFODETGKRNROSGSFLLSTGMFTMIALBELCEIIVVYGWSDSYCR 240  
Db 181 VYTFETEMMAVYCDQIFODETGKRNROSGSFLLSTGMFTMIALBELCEIIVVYGWSDSYCR 240  
QY 241 EKSHPSVYHYFEKGRLEDCOMYLAEQAPRSARHFTTEKAVFSRMAKKRPYPAHPSMR 300  
| 241 EKSHPSVYHYFEKGRLEDCOMYLAEQAPRSARHFTTEKAVFSRMAKKRPYPAHPSMR 300  
Db 241 EKSHPSVYHYFEKGRLEDCOMYLAEQAPRSARHFTTEKAVFSRMAKKRPYPAHPSMR 300  
QY 301 TE 302  
| 301 TE 302  
Db 301 TE 302

## RESULT 4

US-09-895-298-120  
; Sequence 120, Application US/09895298  
; Publication No. US20030078405A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 47 Human Secreted Proteins  
; FILE REFERENCE: P2035P1  
; CURRENT APPLICATION NUMBER: US/09/895,298  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: 09/591,16  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: PCT/US99/29950  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: 60/113,006  
; PRIOR FILING DATE: 1998-12-18  
; PRIOR APPLICATION NUMBER: 60/112,809  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 120  
; LENGTH: 345  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (280)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; US-09-895-298-120

Query Match 99.8%; Score 1622; DB 10; Length 345;  
Best Local Similarity 99.7%; Pred. No. 7.5e-160;  
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVGPPLHFGYS 60  
| 44 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVGPPLHFGYS 103  
Db 44 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVGPPLHFGYS 103  
QY 61 SVDPGRPLVBPSCSAVSSSGMGLSGAGLIDSAECVFRNQAPTVGFADVGQRST 120  
| 104 SVDPGRPLVBPSCSAVSSSGMGLSGAGLIDSAECVFRNQAPTVGFADVGQRST 163  
Db 104 SVDPGRPLVBPSCSAVSSSGMGLSGAGLIDSAECVFRNQAPTVGFADVGQRST 163  
QY 121 LRVVSHSVLLLRNYSHTYQKADTIYMWGQGRHNDRLVLTGTYTTLQLTMYGLO 180  
| 121 LRVVSHSVLLLRNYSHTYQKADTIYMWGQGRHNDRLVLTGTYTTLQLTMYGLO 223  
Db 121 LRVVSHSVLLLRNYSHTYQKADTIYMWGQGRHNDRLVLTGTYTTLQLTMYGLO 223  
QY 181 VYTFETEMMAVYCDQIFODETGKRNROSGSFLLSTGMFTMIALBELCEIIVVYGWSDSYCR 240  
| 224 VYTFETEMMAVYCDQIFODETGKRNROSGSFLLSTGMFTMIALBELCEIIVVYGWSDSYCR 283  
Db 224 VYTFETEMMAVYCDQIFODETGKRNROSGSFLLSTGMFTMIALBELCEIIVVYGWSDSYCR 283  
QY 241 EKSHPSVYHYFEKGRLEDCOMYLAEQAPRSARHFTTEKAVFSRMAKKRPYPAHPSMR 300  
| 284 EKSHPSVYHYFEKGRLEDCOMYLAEQAPRSARHFTTEKAVFSRMAKKRPYPAHPSMR 343  
Db 284 EKSHPSVYHYFEKGRLEDCOMYLAEQAPRSARHFTTEKAVFSRMAKKRPYPAHPSMR 343  
QY 301 TE 302  
| 301 TE 302  
Db 301 TE 302

## RESULT 5

US-10-262-511-184  
; Sequence 184, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Miller, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Yu, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patuturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkete, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Agee, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; PRIOR FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; PRIOR FILING DATE: 2001-10-05  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: Curaseqdist version 0.1  
; SEQ ID NO 184  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-262-511-184

Query Match 71.8%; Score 1167; DB 15; Length 218;  
Best Local Similarity 100.0%; Pred. No. 9.2e-113;  
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 MLAGGLGAEIDSAECVFRNQAPTVGFADVGQRSTLRVVSHTSVPLLRNYSHTYQKAR 144  
| 1 MLAGGLGAEIDSAECVFRNQAPTVGFADVGQRSTLRVVSHTSVPLLRNYSHTYQKAR 60  
Db 1 MLAGGLGAEIDSAECVFRNQAPTVGFADVGQRSTLRVVSHTSVPLLRNYSHTYQKAR 60

QY 145 DTLVWVGGRHMDRVLGRTYRLLQITMYPELOVYTFERMAYCDOIFODETEKNR 204  
 DB 61 DTLVWVGGRHMDRVLGRTYRLLQITMYPELOVYTFERMAYCDOIFODETEKNR 120  
 QY 205 ROSGSPSTGWTMTLALCEELIYVGVMSDYSYCREKSHSPVHYFEKGRLEDCOMYL 264  
 DB 121 ROSGSPSTGWTMTLALCEELIYVGVMSDYSYCREKSHSPVHYFEKGRLEDCOMYL 180  
 QY 265 AHEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSMTE 302  
 DB 181 AHEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSMTE 218

RESULT 6  
 US-10-844-874-20  
 ; Sequence 20, Application US/10844874  
 ; Publication No. US20040204381A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moskal, Joseph  
 ; APPLICANT: Yamamoto, Hirotsuka  
 ; TITLE OF INVENTION: Detection and Treatment of Glyco-Enzyme-Related Disease  
 ; FILE REFERENCE: 97-186-E  
 ; CURRENT APPLICATION NUMBER: US/10/844,874  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 20  
 ; LENGTH: 305  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: rat alpha 2, 6-ST  
 US-10-844-874-20

Query Match 44.7%; Score 726.5; DB 17; Length 305;  
 Best Local Similarity 48.7%; Pred. No. 1, 1e-66;  
 Matches 146; Conservative 46; Mismatches 85; Indels 23; Gaps 5;  
 QY 14 SVFSAVYILLCCMAGPLCLATCLDHFPT-----GSRPTVROPLHFS-----CY 59  
 DB 12 AVSFILACILL-----LMRLANDVTFPLLNCFOQPKTKMPLSYTLRQPLQTHVGY 64  
 QY 60 SSVPDGPVLEPGRSCAVVSSSGQMLGSGIGAEIDAEVCFFRNNOAPTGFEDVQGRS 119  
 DB 65 INVRTOEPL-QLNCHCAVVSNSQMGQVGEIDASCIWNNAPPTGFEEDVGYMT 123  
 QY 120 TLRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRLLQITMYPEL 179  
 DB 124 MVRVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRLLQITMYPEL 183  
 QY 180 QVTFETERMAYCDOIFODETEKNRROSGSPSTGWTMTLALCEELIYVGVMSDYSY 239  
 DB 184 QIVYTFEQRMVYCDGVFKETGKDRVSGSYLSTGWFTEFLADACSHVYGKINLTYC 243  
 QY 240 REKSHSPVHYFEKGRLEDCOMYLAHEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSW 299  
 DB 244 TTGSGYRVPVHYHYEQGR-DECNEYLHHEHA PYGGHRTITTEKVPFAKAKKGRIVFTHPNW 302

RESULT 7  
 US-09-833-245-1914  
 ; Sequence 1914, Application US/09833245  
 ; Publication No. US20040010134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546PCT  
 ; CURRENT APPLICATION NUMBER: US/09/833,245  
 ; CURRENT FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229, 358  
 ; PRIOR FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: 60/256, 931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199, 384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1914  
 ; LENGTH: 305  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-833-245-1914

Query Match 44.1%; Score 716.5; DB 11; Length 305;  
 Best Local Similarity 52.7%; Pred. No. 1, 2e-65;  
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;  
 QY 55 HESGYSVPDGPVLEPGRSCAVVSSSGQMLGSGIGAEIDAEVCFFRNNOAPTGFED 114  
 DB 61 HY-GYINVKTQEPLOLD-CDLCAIVSNSQMGQVGEIDRSSCIWNNAPPTGXEED 118  
 QY 115 VGORSTLRVSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRLLQITR 174  
 DB 119 VGMATMIRVSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRLLQITR 178  
 QY 175 MYPLQVYTFETERMAYCDOIFODETEKNRROSGSPSTGWTMTLALCEELIYVGV 234  
 DB 179 IYVNAQIVYTFETERMAYCDOIFODETEKNRROSGSPSTGWTMTLADACSHVYGM 238  
 QY 235 SDYSYCREKSHSPVHYFEKGRLEDCOMYLAHEQAPRSARHFTTEKAVFSRWAKKRPV 294  
 DB 239 NDTYCKTEGVRKVPVHYHYEQGR-DECNEYFLHHEHAPYGGHRTITTEKVPFAKAKKGR 297  
 QY 295 AHPSPW 299  
 DB 298 THPNW 302

RESULT 8  
 US-09-833-245-1915  
 ; Sequence 1915, Application US/09833245  
 ; Publication No. US20040010134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546PCT  
 ; CURRENT APPLICATION NUMBER: US/09/833,245  
 ; CURRENT FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229, 358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256, 931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199, 384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1915  
 ; LENGTH: 305  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-833-245-1915

Query Match 44.1%; Score 716.5; DB 11; Length 305;  
 Best Local Similarity 52.7%; Pred. No. 1, 2e-65;  
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;  
 QY 55 HESGYSVPDGPVLEPGRSCAVVSSSGQMLGSGIGAEIDAEVCFFRNNOAPTGFED 114  
 DB 61 HY-GYINVKTQEPLOLD-CDLCAIVSNSQMGQVGEIDRSSCIWNNAPPTGXEED 118  
 QY 115 VGORSTLRVSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRLLQITR 174  
 DB 119 VGMATMIRVSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRLLQITR 178



```
Qy 175 MYPELQVYTFTEHMAVCDQIFODETGKNRSGSFLSTGFWTILALCEETIVVYGMV 234
Db 179 IYPAQIYVTEKMSYCDGVFKKETGKDRVQSSYSTIGWFTILAMDACYGHHVGM 238
Qy 235 SDSYCREKSHPSVYHYFEKGRDLDECOMYLAEQAPRSARHPTTEKAVFSRMAKRPV 294
Db 239 NDTYCKTEGVRKRVYHYEQR-DECODEYFLHEHAPYGHRFITEKKVFAKMAKGRH 297
Qy 295 AHPBW 299
Db 298 THPMW 302

RESULT 9
US-10-188-186-22
; Sequence 22, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT FILING DATE: 2002-07-02
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 22
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-22

Query Match 43.9%; Score 714.5; DB 15; Length 305;
Best Local Similarity 52.2%; Pred. No. 1.9e-65;
Matches 128; Conservative 49; Mismatches 65; Indels 3; Gaps 3;

Qy 55 HFGSYSSVPDGKPLVREPCRSQAVSSSGQMLSGLAGAIDSAECVFRMNAQPTVGF 114
Db 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQVGVQVGENEIDRSSCIWRMNNAPTKGYE 118
Qy 115 VGRSTIRVYSHSVPLLNRYSHYFQKARDTLVMVQGGRRHMDRVLGRTYRTLLQ 174
Db 119 VGRMTIRVYSHSVPLLNRYSHYFQKARDTLVMVQGGRRHMDRVLGRTYRTLLQ 178
Qy 175 MYPELQVYTFTEHMAVCDQIFODETGKNRSGSFLSTGFWTILALCEETIVVYGMV 234
Db 179 IYPAQIYVTEKMSYCDGVFKKETGKDRVQSSYSTIGWFTILAMDACYGHHVGM 238
Qy 235 SDSYCREKSHPSVYHYFEKGRDLDECOMYLAEQAPRSARHPTTEKAVFSRMAKRPV 294
Db 239 NDTYCKTEGVRKRVYHYEQR-DECODEYFLHEHAPYGHRFITEKKVFAKMAKGRH 297
Qy 295 AHPBW 299
```

```
Db 298 THPMW 302

RESULT 10
US-10-104-047-2350
; Sequence 2350, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2350
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2350

Query Match 43.9%; Score 713.5; DB 14; Length 305;
Best Local Similarity 52.2%; Pred. No. 2.4e-65;
Matches 128; Conservative 49; Mismatches 65; Indels 3; Gaps 3;

Qy 55 HFGSYSSVPDGKPLVREPCRSQAVSSSGQMLSGLAGAIDSAECVFRMNAQPTVGF 114
Db 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQVGVQVGENEIDRSSCIWRMNNAPTKGYE 118
Qy 115 VGRSTIRVYSHSVPLLNRYSHYFQKARDTLVMVQGGRRHMDRVLGRTYRTLLQ 174
Db 119 VGRMTIRVYSHSVPLLNRYSHYFQKARDTLVMVQGGRRHMDRVLGRTYRTLLQ 178
Qy 175 MYPELQVYTFTEHMAVCDQIFODETGKNRSGSFLSTGFWTILALCEETIVVYGMV 234
Db 179 IYPAQIYVTEKMSYCDGVFKKETGKDRVQSSYSTIGWFTILAMDACYGHHVGM 238
Qy 235 SDSYCREKSHPSVYHYFEKGRDLDECOMYLAEQAPRSARHPTTEKAVFSRMAKRPV 294
Db 239 NDTYCKTEGVRKRVYHYEQR-DECODEYFLHEHAPYGHRFITEKKVFAKMAKGRH 297
Qy 295 AHPBW 299
Db 298 THPMW 302

RESULT 11
US-10-188-186-24
; Sequence 24, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT FILING DATE: 2002-07-02
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
```

APPLICANT: UOMH MONAHAN  
APPLICANT: Manjula GANNANAPARAPU  
APPLICANT: Sebastian HOERSCH  
APPLICANT: ShubhangI KAWATKAR  
APPLICANT: Steve G. KOVATS  
APPLICANT: Rachel E. MEYERS  
APPLICANT: Michael MORRISSEY  
APPLICANT: Peter OLANDT  
APPLICANT: Aml SEN  
APPLICANT: Peter VERIBY  
APPLICANT: Gordon B. MILLS  
APPLICANT: Robert C. BAST, Jr.  
APPLICANT: Karen LU  
APPLICANT: Rosemarie SCHMANDT  
APPLICANT: Xumei ZHAO  
APPLICANT: Karen GLAVITT  
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification  
TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
FILE REFERENCE: MRI-030  
CURRENT APPLICATION NUMBER: US/10/097,340  
CURRENT FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 60/276,025  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/325,149  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/276,026  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/324,967  
PRIOR FILING DATE: 2001/09/26  
PRIOR APPLICATION NUMBER: 60/311,732  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: 60/325,102  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/323,580  
PRIOR FILING DATE: 2001-09-19  
NUMBER OF SEQ ID NOS: 363  
SOFTWARE: FastSeq for Windows Version 4.0



Query Match	39.1%;	Score 636.5;	DB 14;	Length 336;
Best Local Similarity	47.5%;	Pred. No. 2.9e-57;		
Matches 124;	Conservative 45;	Mismatches 89;	Indels 3;	Gaps 3;

```

QY      4 TGRSPFLV-GPLHFGSGVSPDGKPLVPRGSGCAVSSGOMGSLGAIIDAEVCVR 102
Db      64 TQGRPGVPRAGRPLDGLGIVADHKRL-KKHCDCALVSSGHLHSGSGSIDDTCEVTR 122
QY      103 MNQAPLVGEADYGGSRSTRVSVSHTSVPLLNYSHYFOQAKDPLVYVGGRRMDVLG 162
Db      123 MNQAPLVGGYGRDVGNTSRSLVIAHSISORILINRDLNVSGQVFIEMGSSYMRDGGK 182
QY      163 GRTYRTLLDTTRVYPCGLQVYTFTERMMVSCDIFQDETGGKRRGSGSLSTGWTMLAL 222
Db      183 GQYNNHLHLSSQVLPYLKAPMLTRIKMLQFDLTFQEGTGKRRKLSNTWLSGWTMLIAL 242
QY      223 ELCEEIVYGVWSDSYCREKSHSPVYHYFEKGRLLDECOMYLAHQAPR-SAHRRFITEKA 281
Db      243 ELCDRIINVGVMPDPDCRPNHPSVYHYUDDFEGDECTMYLSHERGKSGSHRRFITEKR 302
QY      282 VFSRMAKRRPIVFAHPSWRTS 302
Db      303 VFSKMAETFIHFQPDMAKPE 323

```

Search completed: December 10, 2004, 18:29:16  
Job time : 148 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:03:47 ; Search time 156 Seconds  
(without alignments)  
694.463 Million cell updates/sec

Title: US-10-019-114A-7  
Perfect score: 1626  
Sequence: 1 MKAFGRVLIIICVFEVSAV.....FSRWAKRPPIVFAHPSWRT 302

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: geneeqp1980s:\*  
2: geneeqp1990s:\*  
3: geneeqp2000s:\*  
4: geneeqp2001s:\*  
5: geneeqp2002s:\*  
6: geneeqp2003as:\*  
7: geneeqp2003bs:\*  
8: geneeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1626	100.0	302 3 AAB25764	AAB25764 Human sec
2	1626	100.0	302 3 AAB28674	AAB28674 Human car
3	1626	100.0	302 4 AAB61614	AAB61614 Human pro
4	1626	100.0	302 4 AAB75350	AAB75350 Human sec
5	1626	100.0	302 7 ADD47793	Add47793 Human pro
6	1626	100.0	302 8 ADP19151	ADP19151 Human sec
7	1622	99.8	303 3 AAB24441	AAB24441 Human sec
8	1622	99.8	345 3 AAB24495	AAB24495 Human sec
9	1167	71.8	218 6 AD05824	AD05824 Human NOV
10	1159	71.3	218 6 ADN62987	ADN62987 Human NOV
11	726.5	44.7	305 7 ADD47791	Add47791 Rat Prote
12	716.5	44.1	305 5 AAE21578	AAE21578 Human gen
13	716.5	44.1	305 5 AAE21595	AAE21595 Human gen
14	716.5	44.1	305 5 ABG65165	ABG65165 Human alb
15	716.5	44.1	305 5 ABG65166	ABG65166 Human alb
16	716.5	44.1	305 8 ADL78432	ADL78432 Albumin f
17	716.5	44.1	305 8 ADL78433	ADL78433 Albumin f
18	716.5	44.1	305 5 AAE21617	AAE21617 Human gen
19	714.5	43.9	305 7 ADC26197	ADC26197 Human NOV
20	713.5	43.9	305 7 ADB64196	ADB64196 Human pro
21	694	42.7	318 7 ADC26199	ADC26199 Human NOV
22	677.5	41.7	305 7 ADC26195	ADC26195 Human NOV
23	639.5	39.3	336 4 AAE10705	AAE10705 Human sia
24	639.5	39.3	336 5 ABG96367	ABG96367 Human ova
25	639.5	39.3	336 6 ADA54873	ADA54873 Human pro

26	636.5	39.1	336 5 AAE22149	AAE22149 Human TRN
27	633.5	39.0	275 4 AAM41882	AAM41882 Human pol
28	595.5	36.6	246 4 AAM40096	AAM40096 Human pol
29	548.5	33.7	256 4 AAM41144	AAM41144 Human pol
30	537	33.0	299 3 AAY99356	AAY99356 Human PRO
31	537	33.0	299 4 AAB66105	AAB66105 Protein o
32	537	33.0	299 4 AAU29126	AAU29126 Human PRO
33	537	33.0	299 6 ABUS5802	ABUS5802 Human PRO
34	537	33.0	299 6 ABUS8050	ABUS8050 Novel hum
35	537	33.0	299 6 ABUS8365	ABUS8365 Human sec
36	537	33.0	299 6 ABR66239	ABR66239 Human sec
37	537	33.0	299 6 ABR65629	ABR65629 Human sec
38	537	33.0	299 6 ABUS9569	ABUS9569 Human sec
39	537	33.0	299 6 ABUS2808	ABUS2808 Human PRO
40	537	33.0	299 6 ABUS9929	ABUS9929 Novel hum
41	537	33.0	299 6 ABR68178	ABR68178 Human sec
42	537	33.0	299 6 ABUS96231	ABUS96231 Novel hum
43	537	33.0	299 6 ABUS2662	ABUS2662 Human sec
44	537	33.0	299 6 ABO08739	ABO08739 Human sec
45	537	33.0	299 6 ABO02791	ABO02791 Human sec

## ALIGNMENTS

RESULT 1  
AAB25764  
ID AAB25764 standard; protein; 302 AA.

XX AAB25764;

DT 28-NOV-2000 (first entry) date of database entry  
XX Human secreted protein SEQ ID #76.

XX Human; secreted protein; forensic procedure; gene therapy;  
KW Chromosome mapping; cancer; autoimmune disease; cardiovascular disorder;  
KW Cystic fibrosis; hypothyroidism; immunological disorder; amyloidosis;  
KW brain disorder; skeletal muscle disorder; eye disorder; obesity;  
KW mitochondrial cytopathy; diabetes; atherosclerosis; Alzheimer's disease;  
KW neurodegenerative disorder; graft rejection; dementia; hyperlipidaemia;  
KW septic shock; impotence.

XX Homo sapiens.

XX WO200037491-A2.

PD 29-JUN-2000. *patent publication date / issued*  
PR 20-DEC-1999; 99MC-IB002058. *patent filing date / submitted*  
PR 22-DEC-1998; 98US-0113686P. *priority date*  
PR 25-JUN-1999; 99US-0141032P.

XX (GERT) GENSET.

XX Bougueleret L, Dumas J, Duclert A;  
XX WPI; 2000-442637/38.  
XX DR N-PSDB; AAA87726.  
XX PT Polynucleotides and polypeptides encoding proteins with signal peptides,  
XX PT useful in diagnostic, forensic, gene therapy and chromosome mapping  
XX procedures.

PS Claim 9; Fig 12; 306pp; English.

XX This sequence represents a human secreted protein amino acid sequence.  
XX The invention relates to sequences AAB25763-A87774 which encode human  
XX secreted proteins AAB25763-A85812. The proteins include signal peptides.  
XX Included in the invention are a host cell containing one of the cDNA  
XX sequences, and a purified antibody capable of binding to one of the  
XX secreted proteins. Also contained in the invention are methods for

CC storing the sequence data on a computer system, and a method for  
 CC identifying features of the cDNA sequences using a computer programme.  
 CC The cDNAs are useful for expressing secreted proteins or fragments to  
 CC obtain antibodies capable of specifically binding to the secreted  
 CC proteins. The cDNAs may also be useful in diagnostic, forensic, gene  
 CC therapy and chromosome mapping procedures and may be used to design  
 CC expression vectors and secretion vectors. The proteins of the invention  
 CC may be used to treat diseases including cancer, autoimmune diseases,  
 CC cardiovascular disorders, cystic fibrosis, hypothyroidism, immunological  
 CC disorders, amyloidosis, brain disorders, skeletal muscle disorders, eye  
 CC disorders, obesity, mitochondrial cytopathies, diabetes, atherosclerosis,  
 CC neurodegenerative disorders, graft rejection, Alzheimer's disease,  
 CC dementia, hyperlipidaemia, septic shock and impotence  
 XX  
 XX Sequence 302 AA;  
 SQ  
 Query Match 100.0%; Score 1626; DB 3; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-160;  
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKAPGRVLVLIILCSVFSANVYIILCCWAGLPCLCLATLDDHFFPGSRPTVGPGLHFSGYG 60  
 DB 1 MKAPGRVLVLIILCSVFSANVYIILCCWAGLPCLCLATLDDHFFPGSRPTVGPGLHFSGYG 60  
 QY 61 SVPDGKPLVREPCRSQAVSSSGOMLGSGIAGETDSECVFRNMQAPTVGFEDVGQRST 120  
 DB 61 SVPDGKPLVREPCRSQAVSSSGOMLGSGIAGETDSECVFRNMQAPTVGFEDVGQRST 120  
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGQGRMDRVLGGRYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGQGRMDRVLGGRYRTLLQLTRMYPGLQ 180  
 QY 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTLALCEIIVYGVWVSDSYCR 240  
 DB 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTLALCEIIVYGVWVSDSYCR 240  
 QY 241 EKSHPSVPHYFEKGRGLDECOMYLAHQAPRSARHRTTEKAVSRNAKKRPVFAHPSWR 300  
 DB 241 EKSHPSVPHYFEKGRGLDECOMYLAHQAPRSARHRTTEKAVSRNAKKRPVFAHPSWR 300  
 QY 301 TE 302  
 DB 301 TE 302  
 RESULT 2  
 AAB28674  
 ID AAB28674 standard; protein; 302 AA.  
 XX  
 AC AAB28674;  
 DT 13-FEB-2001 (first entry)  
 DE Human carbohydrate-modifying enzyme Incyte ID No: 983984CD1.  
 XX  
 KW Human; carbohydrate-modifying enzyme; CME; antidiabetic;  
 KW immunosuppressive; anti-HIV; antiinflammatory; antianaemic;  
 KW antidiabetic; antiarteriosclerotic; antilyroid; hepatocytic;  
 KW nephrotic; antilygout; thyromimetic; neuroprotective; osteoparhic;  
 KW antidiabetic; antipsoriasis; uroparhic; ophthalmological;  
 KW dermatological; antilyser; cytotoxic; virucide; antibacterial;  
 KW fungicide; protozoicide; tranquilliser; vulnerary; diabetes;  
 KW autoimmune disorder; inflammatory disorder; infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200063351-A2.  
 XX  
 PD 26-OCT-2000.  
 XX  
 PF 20-APR-2000; 2000WO-USO10882.  
 XX  
 PR 21-APR-1999; 99US-0130383P.

XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Lal P, Yue H, Tang YT, Hillman JL, Baughn MR, Yang J;  
 XX  
 DR WPI; 2000-672729/65.  
 DR N-PSDB; AAC65393.  
 XX  
 PT Novel carbohydrate modifying enzyme polypeptides and polynucleotides for  
 PT diagnosis, treatment, and prevention of carbohydrate metabolism  
 PT disorders, autoimmune/inflammatory disorders, and cancer.  
 PS  
 PS Claim 1; Page 67-68; 75pp; English.  
 XX  
 CC The present sequence is a human carbohydrate-modifying enzyme (CME). CME  
 CC polynucleotides and polypeptides are useful for treating and diagnosing  
 CC diseases associated with CME such as diabetes, autoimmune/inflammatory  
 CC disorders such as AIDS, Addison's disease, adult respiratory distress  
 CC syndrome, allergies, anaemia, asthma, atherosclerosis, autoimmune  
 CC thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's  
 CC disease, emphysema, erythroblastosis fetalis, glomerulonephritis, Good  
 CC pasture's syndrome, gout, Grave's disease, Hashimoto's thyroiditis,  
 CC multiple sclerosis, myasthenia gravis, osteoarthritis, osteoporosis,  
 CC pancreatitis, polymyositis, psoriasis, Reiter's syndrome, arthritis,  
 CC scleroderma, Sjogren's syndrome, systemic lupus erythematosus, ulcerative  
 CC colitis, uveitis, Werner syndrome, complications of cancer,  
 CC haemodialysis, and extracorporeal circulation, viral, bacterial, fungal  
 CC parasitic, protozoal, and helminthic infections, trauma, or cancer. CME,  
 CC or its catalytic or immunogenic fragment, is useful for drug screening  
 XX  
 SQ Sequence 302 AA;  
 Query Match 100.0%; Score 1626; DB 3; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-160;  
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKAPGRVLVLIILCSVFSANVYIILCCWAGLPCLCLATLDDHFFPGSRPTVGPGLHFSGYG 60  
 DB 1 MKAPGRVLVLIILCSVFSANVYIILCCWAGLPCLCLATLDDHFFPGSRPTVGPGLHFSGYG 60  
 QY 61 SVPDGKPLVREPCRSQAVSSSGOMLGSGIAGETDSECVFRNMQAPTVGFEDVGQRST 120  
 DB 61 SVPDGKPLVREPCRSQAVSSSGOMLGSGIAGETDSECVFRNMQAPTVGFEDVGQRST 120  
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGQGRMDRVLGGRYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGQGRMDRVLGGRYRTLLQLTRMYPGLQ 180  
 QY 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTLALCEIIVYGVWVSDSYCR 240  
 DB 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTLALCEIIVYGVWVSDSYCR 240  
 QY 241 EKSHPSVPHYFEKGRGLDECOMYLAHQAPRSARHRTTEKAVSRNAKKRPVFAHPSWR 300  
 DB 241 EKSHPSVPHYFEKGRGLDECOMYLAHQAPRSARHRTTEKAVSRNAKKRPVFAHPSWR 300  
 QY 301 TE 302  
 DB 301 TE 302  
 RESULT 3  
 AAB61614  
 ID AAB61614 standard; protein; 302 AA.  
 XX  
 AC AAB61614;  
 DT 05-APR-2001 (first entry)  
 DE Human protein HP03380.  
 XX  
 KW Human; hydriphobic domain; immune deficiency; autoimmune disorder;  
 KW allergy; tissue growth; regeneration; wound healing; burn; tumour;





XX	ADD47793	
ID	ADD47793	standard; protein; 302 AA.
XX		
AC	ADD47793;	
XX		
DT	29-JAN-2004	(first entry)
XX		
DE	Human Protein CAC07404,	SEQ ID NO 13489.
XX		
KW	Human; pain; neuronal tissue; gene therapy;	
XX	spinal segmental nerve injury; chronic constriction injury; CCI;	
OS	spared nerve injury; SNI; Chung.	
XX		
XX	Homo sapiens.	
XX		
XX	WO2003016475-A2.	
XX		
PD	<u>27-FEB-2003.</u>	
XX		
PF	14-AUG-2002; 2002WO-US025765.	
XX		
PR	14-AUG-2001; 2001US-0312147P.	
XX	01-NOV-2001; 2001US-0346382P.	
PR	26-NOV-2001; 2001US-0333347P.	
XX		
PA	(GHEO ) GEN HOSPITAL CORP.	
PA	(FARB ) BAYER AG.	
PI		
PI	Woolf C, D'urso D, Befort K, Costigan M;	
XX		
DR	WPI; 2003-268312/26.	
DR	GENBANK; CAC07404.	
XX		
PT	New composition comprising two or more isolated polypeptides, useful for	
XX	preparing a medicament for treating pain in an animal.	
PS	Claim 1; Page; 1017pp; English.	
XX		
CC	The invention discloses a composition comprising two or more isolated rath	
CC	or human polynucleotides or a polynucleotide which represents a fragment,	
CC	derivative or allelic variation of the nucleic acid sequence. Also	
CC	claimed are a vector comprising the novel polynucleotide, a host cell	
CC	comprising the vector, a method for identifying a nucleotide sequence	
CC	which is differentially regulated in an animal subjected to pain and a	
CC	kit to perform the method, an array, a method for identifying an agent	
CC	that increases or decreases the expression of the polynucleotide sequence	
CC	that is differentially expressed in neuronal tissue of a first animal	
CC	subjected to pain, a method for identifying a compound which regulates	
CC	the expression of a polynucleotide sequence which is differentially	
CC	expressed in an animal subjected to pain, a method for identifying a	
CC	compound that regulates the activity of one or more of the	
CC	polynucleotides, a method for producing a pharmaceutical composition, a	
CC	method for identifying a compound or small molecule that regulates the	
CC	activity in an animal of one or more of the polypeptides given in the	
CC	specification, a method for identifying a compound useful in treating	
CC	pain and a pharmaceutical composition comprising the one or more	
CC	polypeptides or their antibodies. The polynucleotide or the compound that	
CC	modulates its activity is useful for preparing a medicament for treating	
CC	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction	
CC	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene	
CC	therapy). The sequence presented is a human protein (shown in Table 2 of	
CC	the specification) which is differentially expressed during pain. Note:	
CC	The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic form directly from WIPO at	
CC	ftp://wipo.int/pub/published_pct_sequences.	

Query Match	100.0%	Score 1626	DB 7	length 302
Best Local Similarity	100.0%	Pred. No. 1.2e-160		
Matches 302	Conservative 0	Mismatches 0	Indels 0	Gaps 0
by	1	MKAPRLVILICCSVFSAYIILCCWAGIPLCLATLCTDHHFPTGSRPTVPGDLHFGYS	60	

Db	1	MAAPRLVLIILCSVFSAVYILLCCMAGLPACTLATCLDHPPTGSRPIVPGPIHPSGYS	60
QY	61	SVPOGKPLVRECRCAVAVSSSGOMLGSLGAEIDSAECVFPMMNOAPITGVFENDVGORST	120
Db	61	SVPOGKPLVRECRCAVAVSSSGOMLGSLGAEIDSAECVFPMMNOAPITGVFENDVGORST	120
QY	121	LRVASHTVPLLLRNRYSHYFQKARDTLVYVWQCGRHMDRVLGGRVYRTLLQLTRMPGLQ	180
Db	121	LRVASHTVPLLLRNRYSHYFQKARDTLVYVWQCGRHMDRVLGGRVYRTLLQLTRMPGLQ	180
QY	181	VYTFETERMAAYCDQIFODETGKNNRQSGSFLSTGFTMLALCEIIVYGVMSYSYCR	240
Db	181	VYTFETERMAAYCDQIFODETGKNNRQSGSFLSTGFTMLALCEIIVYGVMSYSYCR	240
QY	241	EKSHSIVYHYHFEKRLDECCOMYLAEQAPRSAPHPITEKAVFSRRAKTRPIYFAHPSWR	300
Db	241	EKSHSIVYHYHFEKRLDECCOMYLAEQAPRSAPHPITEKAVFSRRAKTRPIYFAHPSWR	300
QY	301	TE 302	
Db	301	TE 302	
RESULT 6			
ID	ADP19151		
XX	ADP19151	standard; protein; 302 AA.	
AC	ADP19151;		
XX	DT	26-AUG-2004 (first entry)	
DE		Human secreted polypeptide #2.	
XX		Human; secreted protein; genetic disease.	
XX	OS	Homo sapiens.	
XX	US	US2004110939-A1.	
FN		10-JUN-2004.	
PD		15-OCT-2001; 2001US-00978360.	
XX	PF	17-DEC-1998; 98WO-1B002122.	
PR		09-FEB-1999; 99WO-1B000282.	
PR		21-JUN-2000; 2000WO-1B000951.	
PR		15-SEP-2000; 2000US-00663600.	
XX		(GEST ) GENSET SA.	
PA		Dumas Milne Edwards J, Bougueleret L, Jobert S, Clusel C;	
XX	PI	Duchert A;	
XX	DR	WPI; 2004-440404/41.	
XX	DR	N-PESDB; ADP18746.	
PT		New isolated polynucleotide encoding secreted polypeptide, useful for	
PT		gene therapy, or in diagnostic procedures to identify individuals having	
XX		genetic diseases resulting from abnormal expression of the genes.	
XX		Claim 2; SEQ ID NO 407; 113pp; English.	
XX		The invention relates to human cDNA sequences that encode human secreted	
CC		proteins. The invention also relates to an antibody that specifically	
CC		binds to a polypeptide of the invention and a method of binding the	
CC		polypeptide to an antibody. The polynucleotides are useful for expressing	
CC		the entire secreted proteins which they encode and for distinguishing	
CC		human tissues and cells from non-human tissues and cells, and for	
CC		distinguishing between human tissues and cells that do or do not express	
CC		the polynucleotides comprising the cDNAs. The polynucleotides and	
CC		polypeptides are useful in forensic procedures or diagnostic procedures	
CC		to identify individuals with genetic diseases resulting from abnormal	

expression of the genes corresponding to the cDNAs. The sequences are also useful in gene therapy to control or treat genetic diseases. This sequence represents a human secreted polypeptide of the invention. Note: The sequence data for this patent did not form part of the printed CC specification but was obtained in electronic format from USPTO at CC seqdata.uspto.gov/sequence.html.

XX Sequence 302 AA:

Query Match 100.0%; Score 1626; DB 8; Length 302;  
Best Local Similarity 100.0%; Pred. No. 1.2e-160;  
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLVLIILCSVFSAYIILCCWAGPLCLATCTLDHHPFGSPPTVPGPLHFGYS 60  
DB 1 MKAPGRVLVLIILCSVFSAYIILCCWAGPLCLATCTLDHHPFGSPPTVPGPLHFGYS 60  
QY 61 SVPDGKPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVQGRST 120  
DB 61 SVPDGKPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVQGRST 120  
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
QY 181 VYFTFERMAVYCDQIFODETGKRRROSGSLSTGFMITLALCEIIVVYGWVSDSYCR 240  
DB 181 VYFTFERMAVYCDQIFODETGKRRROSGSLSTGFMITLALCEIIVVYGWVSDSYCR 240  
QY 241 EKSHPSVPYHYFEKGRLEDECOMYLAEQAPRSARHPTTEKAVFSRMKKRPVFAHPSMR 300  
DB 241 EKSHPSVPYHYFEKGRLEDECOMYLAEQAPRSARHPTTEKAVFSRMKKRPVFAHPSMR 300  
QY 301 TE 302  
DB 301 TE 302

RESULT 7

AAB24441 ID AAB24441 standard; protein; 303 AA.

XX AAB24441;

XX 20-NOV-2000 (first entry)

XX Human secreted protein sequence encoded by gene 5 SEQ ID NO:66.

XX Human; secreted protein; cytosolic; antianaemic; antidiabetic;  
XX antiinflammatory; ophthalmological; antirheumatic; antiarthritic;  
XX antiisoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;  
XX neuroprotective; antimicrobial; antiparkinsonian; cancer;  
XX immune system disorder; angiogenesis; hyperproliferative disorder;  
XX cardiovascular disorder; apoptosis; neurological disease;  
XX infectious disease; wound healing; chromosome 9.

XX Homo sapiens.

XX MO200035937-A1.

XX 22-JUN-2000.

XX 16-DEC-1999; 99WO-05029950.

XX 17-DEC-1998; 98US-0112809P.

XX 18-DEC-1998; 98US-0113006P.

XX (HUMA-) HUMAN GENOME SCT INC.

XX Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J,  
PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K,  
XX WPI, 2000-431566/37.

DR N-PSDB; AAA78385.

XX Forty seven human nucleic acids encoding secreted proteins, useful in the  
PT treatment, prevention and diagnosis of cancers, disorders of the immune  
PT system, angiogenesis disorders, neurological diseases and  
PT hyperproliferative disorders.

XX Claim 11; Page 482-483; 562pp; English.

XX The polynucleotide sequence given in AAA78381 to AAA78432 encode the  
CC human secreted proteins given in AAB24437 to AAB24604. Human secreted  
CC proteins have activities based on the tissues and cells the genes are  
CC expressed in. Examples of activities include: cytosolic; antianaemic;  
CC antidiabetic; antiinflammatory; ophthalmological; antirheumatic;  
CC antiarthritic; antiparkinsonian; antiparkinsonian; anti-HIV;  
CC nootropic; neuroprotective; antimicrobial; antiparkinsonian. Human  
CC secreted protein polynucleotides, polypeptides, and/or diagnosing other  
CC agonists may be useful in treating, preventing, and/or diagnosing other  
CC diseases, disorders, and/or conditions such as: (a) cancers; (b)  
CC disorders of the immune system; (c) angiogenesis disorders; (d)  
CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases  
CC associated with increase apoptosis; (g) neurological diseases; and (h)  
CC infectious diseases. They are also used to promote wound healing.  
CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the  
CC exemplification of the present invention

XX Sequence 303 AA:

Query Match 99.8%; Score 1622; DB 3; Length 303;  
Best Local Similarity 99.7%; Pred. No. 3.3e-160;  
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLVLIILCSVFSAYIILCCWAGPLCLATCTLDHHPFGSPPTVPGPLHFGYS 60  
DB 1 MKAPGRVLVLIILCSVFSAYIILCCWAGPLCLATCTLDHHPFGSPPTVPGPLHFGYS 60  
QY 61 SVPDGKPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVQGRST 120  
DB 61 SVPDGKPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVQGRST 120  
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
QY 181 VYFTFERMAVYCDQIFODETGKRRROSGSLSTGFMITLALCEIIVVYGWVSDSYCR 240  
DB 181 VYFTFERMAVYCDQIFODETGKRRROSGSLSTGFMITLALCEIIVVYGWVSDSYCR 240  
QY 241 EKSHPSVPYHYFEKGRLEDECOMYLAEQAPRSARHPTTEKAVFSRMKKRPVFAHPSMR 300  
DB 241 EKSHPSVPYHYFEKGRLEDECOMYLAEQAPRSARHPTTEKAVFSRMKKRPVFAHPSMR 300  
QY 301 TE 302  
DB 301 TE 302

RESULT 8

AAB24495 ID AAB24495 standard; protein; 345 AA.

XX AAB24495;

XX 20-NOV-2000 (first entry)

XX Human secreted protein sequence encoded by gene 5 SEQ ID NO:120.

XX Human; secreted protein; cytosolic; antianaemic; antidiabetic;  
XX antiinflammatory; ophthalmological; antirheumatic; antiarthritic;  
XX antiisoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;  
XX neuroprotective; antimicrobial; antiparkinsonian; cancer;  
XX immune system disorder; angiogenesis; hyperproliferative disorder;  
XX cardiovascular disorder; apoptosis; neurological disease;

KW Infectious disease; wound healing; chromosome 9.  
 XX Homo sapiens.  
 OS  
 XX WO200035937-A1.  
 XX  
 XX 22-JUN-2000.  
 XX  
 XX 16-DEC-1999; 99WO-US029950.  
 XX  
 XX 17-DEC-1998; 98US-0112809P.  
 XX 18-DEC-1998; 98US-0113006P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;  
 PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;  
 XX WPI; 2000-431566/37.  
 XX  
 XX Forty seven human nucleic acids encoding secreted proteins, useful in the  
 PT treatment, prevention and diagnosis of cancers, disorders of the immune  
 PT system, angiogenesis disorders, neurological diseases and  
 PT hyperproliferative disorders.  
 XX  
 XX Disclosure; Page 19; 562pp; English.  
 XX  
 XX The polynucleotide sequence given in AAA78381 to AAA78432 encode the  
 CC human secreted proteins given in AAB24437 to AAB24604. Human secreted  
 CC proteins have activities based on the tissues and cells the genes are  
 CC expressed in. Examples of activities include: cytostatic; antineoplastic;  
 CC antidiabetic; antiinflammatory; ophthalmological; antirheumatic;  
 CC antithrombotic; antiproliferative; angiogenic; cardiac; anti-HIV;  
 CC neurotrophic; neuroprotective; antimicrobial and antiparkinsonian. Human  
 CC secreted protein polynucleotides, polypeptides, antagonists and/or  
 CC agonists may be useful in treating, preventing, and/or diagnosing other  
 CC diseases, disorders, and/or conditions such as: (a) cancers; (b)  
 CC disorders of the immune system; (c) angiogenesis disorders; (d)  
 CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases  
 CC associated with increase apoptosis; (g) neurological diseases; and (h) diseases  
 CC infectious diseases. They are also used to promote wound healing.  
 CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the  
 CC exemplification of the present invention  
 CC  
 XX Sequence 345 AA;  
 SO  
 Query Match 99.8%; Score 1622; DB 3; Length 345;  
 Best Local Similarity 99.7%; Pred. No. 3.9e-160;  
 Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MKAPGRVLLILCSVPSAVYILLCCVAGPLCLATCLDHHFPTGSRPTVPGPLAFSGYS 60  
 DB 44 MKAPGRVLLILCSVPSAVYILLCCVAGPLCLATCLDHHFPTGSRPTVPGPLAFSGYS 103  
 QY 61 SVDPGRVLPVPCSCAVSSSSGQMGSGIAGIIDSACVFRNQAPTVGFADVGQRST 120  
 DB 104 SVDPGRVLPVPCSCAVSSSSGQMGSGIAGIIDSACVFRNQAPTVGFADVGQRST 163  
 QY 121 LRVVSHSVPLLLRNYSHYFOKARDLTYMWGQGRHMDRLVGRRTYTLQLTRMPGLQ 180  
 DB 164 LRVVSHSVPLLLRNYSHYFOKARDLTYMWGQGRHMDRLVGRRTYTLQLTRMPGLQ 223  
 QY 181 VYFETERMAVYCDQIFODETGKRRRSGSPSTGFMILALEICEEIVVYGWVSDYCR 240  
 DB 224 VYFETERMAVYCDQIFODETGKRRRSGSPSTGFMILALEICEEIVVYGWVSDYCR 283  
 QY 241 EKSHPSVYHYFEKGRLLDECQMTLAHQAPRSARHRTTEKAVPSRMKKRPPIVFAHPSMR 300  
 DB 284 EKSHPSVYHYFEKGRLLDECQMTLAHQAPRSARHRTTEKAVPSRMKKRPPIVFAHPSMR 343  
 QY 301 TE 302  
 DB 344 TE 345

RESULT 9  
 ADA05824  
 ID ADA05824 standard; protein; 218 AA.  
 XX  
 XX ADA05824;  
 AC  
 XX  
 XX 06-NOV-2003 (first entry)  
 DT  
 XX  
 XX Human NOV44a protein SEQ ID NO:184.  
 DE  
 XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;  
 XX immunomodulator; cytostatic; neurotrophic; neuroprotective;  
 KW antiparkinsonian; antidiabetic; gene therapy; human disease;  
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; hematopoietic disorder; dyslipidemia.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO2003029424-A2.  
 PN  
 XX 10-APR-2003.  
 PD  
 XX  
 XX 02-OCT-2002; 2002WO-US031373.  
 PF  
 XX  
 XX 02-OCT-2001; 2001US-0326483P.  
 PR 05-OCT-2001; 2001US-0327435P.  
 PR 05-OCT-2001; 2001US-0327449P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328044P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 12-OCT-2001; 2001US-0329414P.  
 PR 15-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0495755P.  
 PR 01-NOV-2001; 2001US-046357P.  
 PR 17-APR-2002; 2002US-0373260P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373844P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 17-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381542P.  
 PR 28-MAY-2002; 2002US-0383656P.  
 PR 29-MAY-2002; 2002US-0383631P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 PR 01-OCT-2002; 2002US-00262511.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 XX Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ott T, Gorman L, Zehnen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiIppio VA;  
 PI Eisen AJ, Gangoli EA, Rieger DK, Spaderina SK;  
 XX  
 XX WPI; 2003-381626/36.  
 DR N-PSDB; ADA05823.  
 DR  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX Claim 1, Page 268, 586pp; English.

CC The present invention describes NOVX proteins, where X can be 1 to 55  
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
 CC described above and a carrier; (2) a kit comprising, in one or more  
 CC containers, the composition described above; (3) an isolated nucleic acid  
 CC molecule which encodes a NOVX protein of the invention; (4) a vector  
 CC comprising the nucleic acid molecule described above; (5) a cell  
 CC comprising the above vector; (6) an antibody that immunospecifically  
 CC binds to the polypeptide described above; (7) methods for determining the  
 CC presence or amount of the above polypeptide or nucleic acid molecule in a  
 CC sample; (8) methods for determining the presence of or predisposition to  
 CC a disease associated with altered levels of expression of the above  
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
 CC method of identifying an agent that binds to the polypeptide described  
 CC above; (10) a method for identifying a potential therapeutic agent for  
 CC use in treating a pathology that is related to an aberrant expression or  
 CC aberrant physiological interactions of the polypeptide; (11) a method of  
 CC screening for a modulator of activity or of latency or predisposition to  
 CC a pathology associated with the polypeptide; (12) a method for modulating  
 CC the activity of the polypeptide described above; (13) methods of treating  
 CC or preventing a pathology associated with the above polypeptide in a  
 CC mammal; and (14) a method for producing the above polypeptide. NOVX  
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
 CC and antihypertensive activities, and can be used in gene therapy. The  
 CC polypeptide is useful in manufacturing a medicament for treating a  
 CC syndrome associated with a human disease. The polypeptide or the nucleic  
 CC acid molecule may be used to diagnose, treat or prevent metabolic  
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
 CC disease, immune disorders, haematopoietic disorders and various  
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
 CC probes, in chromosome mapping, tissue typing, preventive medicine and  
 CC pharmacogenomics. The present sequence represents a human NOVX from the  
 CC present invention.

XX Sequence 218 AA;

Query Match 71.8%; Score 1167; DB 6; Length 218;

Best Local Similarity 100.0%; Pred. No. 6,1e-113;

Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 MMSGGLAEIDSAECVRNQAFTVGFADVGQSTLRVSHTSVPLLRNYSHYFQKAR 144  
 Db 1 MMSGGLAEIDSAECVRNQAFTVGFADVGQSTLRVSHTSVPLLRNYSHYFQKAR 60  
 QY 145 DTLVWVGQGRHMDRVLCGRYRTLLQITRMYPGLQYTFPERMAYCDQIFODETKNR 204  
 Db 61 DTLVWVGQGRHMDRVLCGRYRTLLQITRMYPGLQYTFPERMAYCDQIFODETKNR 120  
 QY 205 ROSSGFSTMGFTMTLLALELCEELVYVGMVSDSYCRKSHPSVRYHFEKGRLEDCOMYL 264  
 Db 121 ROSSGFSTMGFTMTLLALELCEELVYVGMVSDSYCRKSHPSVRYHFEKGRLEDCOMYL 180  
 QY 265 AHEQAPRSARHFTFEKAVFSRWAKKRPVFAHPSWRT 302  
 Db 181 AHEQAPRSARHFTFEKAVFSRWAKKRPVFAHPSWRT 218

RESULT 10

ID ADN62987 standard; protein, 218 AA.

AC ADN62987;

XX 01-JUL-2004 (first entry)

XX Human NOV44a.

XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;

KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW hemetopietic disorder; dyslipidaemia; metabolic syndrome X;  
 KW wasting disorder.

OS Homo sapiens.

PN US2004038223-A1.

PD 26-FEB-2004.

XX 01-OCT-2002; 2002US-00262511.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 09-OCT-2001; 2001US-0327449P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328849P.

XX 15-OCT-2001; 2001US-0329414P.

XX 17-OCT-2001; 2001US-0330142P.

XX 18-OCT-2001; 2001US-0330309P.

XX 22-OCT-2001; 2001US-0341058P.

XX 24-OCT-2001; 2001US-0342666P.

XX 24-OCT-2001; 2001US-0343629P.

XX 29-OCT-2001; 2001US-0349575P.

XX 01-NOV-2001; 2001US-0346357P.

XX 17-APR-2002; 2002US-0373607P.

XX 19-APR-2002; 2002US-0373815P.

XX 19-APR-2002; 2002US-0373817P.

XX 19-APR-2002; 2002US-0373826P.

XX 19-APR-2002; 2002US-0373844P.

XX 22-APR-2002; 2002US-0374977P.

XX 16-MAY-2002; 2002US-0381037P.

XX 16-MAY-2002; 2002US-0381038P.

XX 16-MAY-2002; 2002US-0381042P.

XX 17-MAY-2002; 2002US-0381642P.

XX 28-MAY-2002; 2002US-0383656P.

XX 29-MAY-2002; 2002US-0383811P.

XX 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J A.

PA (KEKU/) KEKUDA R.

PA (JUUJ/) JU J.

PA (LILL/) LI L.

PA (GUOX/) GUO X.

PA (PATY/) PATTURAJAN M.

PA (SPYT/) SPYTEK K A.

PA (EDIN/) EDINGER S R.

PA (ELLE/) ELLERMAN K.

PA (MALY/) MALYANKAR U M.

PA (ORTT/) ORT T.

PA (GORM/) GORMAN L.

PA (ZERR/) ZERHUSEN B D.

PA (ANDR/) ANDERSON D W.

PA (ZHON/) ZHONG M.

PA (CATT/) CATTERTON E.

PA (JTW/) JI W.

PA (MILL/) MILLER C E.

PA (RAST/) RASTELT L.

PA (STON/) STONE D J.

PA (PENA/) PENNA C B A.

PA (SHEN/) SHENOY S G.

PA (SHIM/) SHIMKETS R A.

PA (ROTH/) ROTHENBERG M E.

PA (LEAC/) LEACH M D.

PA (AGEE/) AGEER M L.

PA (BERG/) BERGHS C.

PA (DIP/) DIPIPPO V A.



DB 12 AVSFIALCIL-----LAMRLANDVTPFLNLNCFGQPKTKM1PLSYTLRQPLQTHYGY 64

QY 60 SSVPDCKPLVREPCRCACAVSSGQMLGSLGAEIDSACVFRMNAQPTVGFADYDQNS 119

DB 65 INRTQEPFL-QLNCHNCACAVSSGQMLGSLGAEIDSACVFRMNAQPTVGFADYDQNS 123

QY 120 TLAVHTSPTLLRLRYSHYFOKARDPLVWVGQGRHMDVAGRTYRTLLQTLTRYPGI 179

DB 124 MVRVSHTSVPLLLKNDPDPFFKASSTTVYVWGFPRMNRKQNGGIYNNMLKKTVDAYPDA 183

QY 180 QVYTFERMAYCDQIFODETGKRRQSGSFLSTGWFMTLALTELCEIIVVYGWVSDSYC 239

DB 184 QIVVTTEQRTYCDGDFKDETKGRVQSGSYLSTGWFMTFLAMDACYSIHVGMINETYC 243

QY 240 REKSHPSVPYHYEKGRLDECOMYLAEQAPRSARHPTTKAVPSRMKKRPVFAHPSW 299

DB 244 TTEGVHAKVPYHYEKGRLDECOMYLAEQAPRSARHPTTKAVPSRMKKRPVFAHPSW 302

RESULT 12

AAE21578

ID AAE21578 standard; protein; 305 AA.

AC AAE21578;

XX 16-JUL-2002 (first entry)

DT Human gene 1 encoded secreted protein HE9PR39, SEQ ID NO:50.

DE Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;

XX foetal abnormality; developmental; haematopoietic disorder; kidney; skin;

XX immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;

XX pregnancy-related disorder; cognitive; neurological; autoimmune disease;

XX inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;

KM Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;

KM diabetes; infection; wound healing; vulnery; chemotaxis; food additive;

KM gene therapy.

XX Homo sapiens.

OS

XX Key

XX Location/Qualifiers

XX Peptide 1..30

XX /label= Signal\_peptide

XX Protein 31..305

XX /label= Mature\_secreted\_protein

XX MO200222654-A1.

XX 21-MAR-2002.

XX 17-JAN-2001; 2001MO-US001385.

XX 12-SEP-2000; 2000US-0231969P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;

XX Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;

XX Fiacella M, Ni J;

XX WPI; 2002-3155684/35.

XX N-PSDB; AAD34106.

XX Seventeen nucleic acid molecules encoding human secreted proteins, useful

XX for treating and preventing cancer, immune disorders (e.g. Addison's

XX disease, and allergies), and cardiovascular disorders (e.g. myocardial

XX ischaemia).

XX Claim 11; Page 437-438; 483pp; English.

XX AAD34106-AA034144 represent cDNAs corresponding to 17 human secreted

XX protein genes, and AAE21578-AAE21616 represent the proteins they encode.

XX AAE21617-AAE21627 represent human secreted protein fragments. The genes

CC and their corresponding secreted proteins are useful for preventing,

CC treating or ameliorating medical conditions, e.g., by protein or gene

CC therapy. Pathological conditions can be diagnosed by determining the presence

CC amount of the new protein in a sample or by determining the presence of

CC mutations in the new genes. Specific uses are described for each of the

CC 17 genes, based on the tissues in which they are most highly expressed,

CC and include developing products for the diagnosis or treatment of cancer,

CC proliferative disorders, tumours, foetal and developmental abnormalities,

CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune

CC diseases (rheumatoid arthritis), inflammation, allergies, neurological

CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,

CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,

CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney

CC disorders, gastrointestinal disorders, pregnancy-related disorders,

CC endocrine disorders, and infections. The proteins can also be used to aid

CC wound healing and epithelial cell proliferation, to prevent skin aging

CC due to sunburn, to maintain organs before transplantation, for supporting

CC cell culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used

CC as a food additive or preservative to modify storage properties. The

CC present sequence represents a human secreted protein of the invention

XX

XX Sequence 305 AA;

XX

Query Match 44.1%; Score 716.5; DB 5; Length 305;

Best Local Similarity 52.7%; Pred. No. 16-65;

Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;

QY 55 HPSGVSSVPDCKPLVREPCRCACAVSSGQMLGSLGAEIDSACVFRMNAQPTVGFADYDQNS 114

DB 61 HY-GYINVTQKQBLQD-CDLCAIVNSGQWQKQVNEIDRSSCIYRMNNAPTKGYEED 118

QY 115 VQGRSLRYVSHRSPVLLRLNRYSHYFOKARDPLVWVGQGRHMDVAGRTYRTLLQTLR 174

DB 119 VGMETIRVSHRSPVLLRLNRYSHYFOKARDPLVWVGQGRHMDVAGRTYRTLLQTLR 178

QY 175 MYPGLQVYTFERMAYCDQIFODETGKRRQSGSFLSTGWFMTLALTELCEIIVVYGWV 234

DB 179 IYVNAQIVYTTTEKGRVQSGSYLSTGWFMTFLAMDACYSIHVGMINETYC 238

QY 235 SDSYCREKSHPSVPYHYEKGRLDECOMYLAEQAPRSARHPTTKAVPSRMKKRPVFAHPSW 294

DB 239 NDTYCTEGYKRPVYHYEKGRLDECOMYLAEQAPRSARHPTTKAVPSRMKKRPVFAHPSW 297

QY 295 AHPSW 299

DB 298 THPNW 302

RESULT 13

AAE21595

ID AAE21595 standard; protein; 305 AA.

XX AAE21595;

XX 16-JUL-2002 (first entry)

DT Human gene 1 encoded secreted protein HE9PR39, SEQ ID NO:67.

XX Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;

XX foetal abnormality; developmental; haematopoietic disorder; kidney; skin;

XX immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;

XX pregnancy-related disorder; cognitive; neurological; autoimmune disease;

XX inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;

KM Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;

KM diabetes; infection; wound healing; vulnery; chemotaxis; food additive;

KM gene therapy.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

XX Peptide 1..30

XX /label= Signal\_peptide





QY 115 VGORSTLRVSVHTSPVLLRNYSHYFOKARDLLYMWGGRHMDRVLGRTYRTLLQLTR 174  
 DB 119 VGRMTMIRVSVHTSPVLLRNYPDFKFKANTTIYIVWGPFRNNRKDNGIYVNNLKKTVG 178  
 QY 175 MYRGLQVYTTTERMMAYCDQIFODEGKNNRQSGSFSTGWMFTMLALEICEBIIVYGMV 234  
 DB 179 IYVNAQIYVTTTERKMSYCDGVFKKETGKDRVQSGSYSTGWMFTFLAMDACYGIHYVGM 238  
 QY 235 SDGYCREKSHPSVYHYHFEKGRDDECOMYLAEQAPSAHRFTTEKAVFSRMKKRPVIF 294  
 DB 239 NDITYCKTEGKRYKPYHYHEQGR-DECDYFLHEHAPYGGHRTTEKKVFAKMAKKRIIF 297  
 QY 295 AHPSW 299  
 DB 298 THPWW 302

RESULT 15  
 ABG65166  
 ID ABG65166 standard; protein: 305 AA.  
 AC ABG65166;  
 XX 27-AUG-2002 (first entry)  
 DT XX  
 DE Human albumin fusion protein #1841.  
 XX  
 KM Albumin fusion protein; therapeutic protein X; human albumin; HA;  
 KM human serum albumin; HSA; cancer; reproductive disorder;  
 KM digestive disorder; immune disorder; endocrine disorder;  
 KM haematopoietic disorder; neural disorder; connective disorder;  
 KM cytostatic; antifertility; antiinflammatory; antiulcer;  
 KM immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;  
 KM neuroprotective; antiParkinsonian; antimicrobial; neuroleptic;  
 KM osteopathic; antiarthritic.  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN W0200177137-A1.  
 PD 18-OCT-2001.  
 XX  
 PF 12-APR-2001; 2001WO-US011988.  
 XX  
 PR 12-APR-2000; 2000US-0229358P.  
 PR 25-APR-2000; 2000US-0199384P.  
 PR 21-DEC-2000; 2000US-0256931P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Haseltine WA;  
 XX  
 DR WPI; 2002-010886/01.  
 XX  
 PT New fusion protein for treating disease e.g. diabetes comprises an  
 albumin fused to a therapeutic protein.  
 XX  
 PS Claim 1; Page 1797-1798; 2102pp; English.  
 XX  
 CC The present invention relates to albumin fusion proteins comprising a  
 CC therapeutic protein X and human albumin (HA, also known as human serum  
 CC albumin, HSA). The proteins are useful for treating a disease or disorder  
 CC that may be modulated by therapeutic protein X. The albumin extends the  
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo  
 CC activity. The protein is useful for treating and diagnosing disorders  
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's  
 CC disease, ulcerative colitis), immune disorders (e.g. acquired  
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),  
 CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,  
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,  
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).  
 CC ABG63326-ABG65518 represent albumin fusion proteins of the invention

XX  
 SQ Sequence 305 AA;  
 Query Match 44.1%; Score 716.5; DB 5; Length 305;  
 Best Local Similarity 52.7%; Pred. No. 1e-65;  
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;

QY 55 HFGSGSVDPGKPLVREPCRSACAVSSGQMLGSGAGAEIDSACVFRNQAFTVGFAD 114  
 DB 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQWQKNGNEIDRSSCIWRNNAPTQGYED 118  
 QY 115 VGORSTLRVSVHTSPVLLRNYSHYFOKARDLLYMWGGRHMDRVLGRTYRTLLQLTR 174  
 DB 119 VGRMTMIRVSVHTSPVLLRNYPDFKFKANTTIYIVWGPFRNNRKDNGIYVNNLKKTVG 178  
 QY 175 MYRGLQVYTTTERMMAYCDQIFODEGKNNRQSGSFSTGWMFTMLALEICEBIIVYGMV 234  
 DB 179 IYVNAQIYVTTTERKMSYCDGVFKKETGKDRVQSGSYSTGWMFTFLAMDACYGIHYVGM 238  
 QY 235 SDGYCREKSHPSVYHYHFEKGRDDECOMYLAEQAPSAHRFTTEKAVFSRMKKRPVIF 294  
 DB 239 NDITYCKTEGKRYKPYHYHEQGR-DECDYFLHEHAPYGGHRTTEKKVFAKMAKKRIIF 297  
 QY 295 AHPSW 299  
 DB 298 THPWW 302

Search completed: December 10, 2004, 18:21:55  
 Job time : 159 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:04:28 ; Search time 192 Seconds

(Without alignments)  
905.016 Million cell updates/sec

Title: US-10-019-114A-7

Perfect score: 1626  
Sequence: 1 MKAPGRVLIIICSVFSAV.....FSRWAKKRPIVFAHPMSWRT 302

Scoring table: BLOSUM62  
Gapop 10.0 , Gapect 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: UniProt 02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1626	100.0	302	1	09441 h alpha-n-a
2	1597	98.2	302	2	062X21
3	1597	98.2	302	2	CAG26699
4	1479	91.0	302	2	Q704X3
5	1479	91.0	302	2	Q704X3
6	1476	90.8	323	2	Q704X4
7	1476	90.8	323	2	Q704X4
8	1452	89.3	302	2	Q6C3J2
9	1452	89.3	317	2	Q6PHP9
10	1452	89.3	317	2	AAH56451
11	1439	88.5	360	1	SI7D MOUSE
12	1436	88.3	302	2	Q6ZX20
13	1436	88.3	302	2	CAG26700
14	1047.5	64.4	289	2	Q704X2
15	1047.5	64.4	289	2	Q704X2
16	740	45.5	306	2	Q704S4
17	740	45.5	306	2	Q704S4
18	740	45.5	306	2	Q704S4
19	733.5	45.1	305	1	SI7C MOUSE
20	733.5	45.1	305	2	BAC28836
21	728.5	44.8	305	2	Q6IN13
22	728.5	44.8	305	2	AAH75501
23	726.5	44.7	305	1	SI7C RAT
24	726.5	44.7	299	2	Q6ZYF0
25	725.5	44.6	299	2	CAG25677
26	724	44.5	292	2	Q6ZX22
27	724	44.5	292	2	CAG26698
28	718.5	44.2	300	2	Q6ZIN9
29	718.5	44.2	300	2	CAG25678
30	717	44.1	234	2	Q9W6U6
31	716.5	44.1	305	2	Q6PC60

32	716.5	44.1	305	2	AAH59363	AAH59363 homo sapi
33	715.5	44.0	305	2	Q6ZYF1	Q6ZYF1 pan troglod
34	715.5	44.0	305	2	CAG25676	CAG25676 pan troglod
35	714.5	43.9	305	1	SI7C HUMAN	Q8NDV1 homo sapien
36	682	41.9	138	2	Q9JHF2	Q9JHF2 mus muscu
37	652	40.1	311	2	Q6ZXV7	Q6ZXV7 brachydanio
38	652	40.1	311	2	CAG26703	CAG26703 brachydan
39	639.5	39.3	335	2	Q6ZXV6	Q6ZXV6 pan troglod
40	639.5	39.3	335	2	CAG26704	CAG26704 pan troglod
41	639.5	39.3	336	1	SI7E HUMAN	Q9BVH7 homo sapien
42	639.5	39.3	336	2	CAD45372	CAD45372 homo sapi
43	630	38.7	335	2	Q8CAM7	Q8CAM7 mus muscu
44	629	38.7	335	2	AAH55737	AAH55737 mus muscu
45	629	38.7	335	2	BAC31331	BAC31331 m 7 days

## ALIGNMENTS

RESULT 1  
SI7D HUMAN STANDARD; PRT; 302 AA.  
ID Q9H4F1; Q9NMU6; Q9UKU1; Q9ULB9; Q9Y3G3; Q9Y3G4;  
AC Q9H4F1; Q9NMU6; Q9UKU1; Q9ULB9; Q9Y3G3; Q9Y3G4;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 05-JUN-2004 (Rel. 44, Last annotation update)  
DE Alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3-N-acetyl-  
DE galactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.7) (NeuAc-  
DE alpha-2,3-Gal-beta-1,3-GalNAc-alpha-2,6-sialyltransferase) (St6GalNAc  
DE IV) (Sialyltransferase 7D) (Sialyltransferase 3C).  
GN Name=SIAT7D; Synonyms=SIAT3C;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=20517254; PubMed=11062056;  
RX Harduin-lepers A., Stokes D.C., Steelant W.F.A., Samyn-Petit B.,  
RA Krzewinski-Rechci M.A., Vallejo-Ruiz V., Zanetta J.P., Auge C.,  
RA Delannoy P.;  
RT "Cloning, expression and gene organization of a human Neusac alpha-2-  
RT Biochem. J. 352:337-348(2000). November issue  
RL [2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Fetal liver;  
RC Kim K.-W., Kim K.-S., Do S.-I., Kim C.-H., Lee Y.-C.;  
RA "Molecular cloning of Neucalalpha2,3galbeta1,3GalNAc alpha2,6-  
RA sialyltransferase cDNA from human fetal liver";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=lung;  
RA Yoshida A.;  
RL "N-acetyl-galactosaminide alpha2,6-sialyltransferase";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX PubMed=14702039; DOI=10.1038/ng1285;  
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M.,  
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Puruya T., Kikkawa E.,  
RA Omura Y., Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M.,  
RA Yamazaki M., Nishimura K., Iwihashi T., Yamashita H., Murakawa K.,  
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,  
RA Ishiguro S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houchu T.,  
RA Kusano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,



DR GO: 001675; F:transferase activity, transferring glycosyl. . .; IEA.  
 DR InterPro: IPR001675; Glyco\_trans\_29.  
 DR Pfam: PF00777; Glyco\_transf\_29; 1.  
 KW Glycosyltransferase; transferase.  
 FT NON\_TER 1  
 FT NON\_TER 302 302  
 SQ SEQUENCE 302 AA; 34138 MW; AB7C984ABE53B553 CRC64;

Query Match 98.2%; Score 1597; DB 2; Length 302;  
 Best Local Similarity 98.0%; Pred. No. 5.5e-137;  
 Matches 296; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVPSAVYIILCCWAGLPCLATCLDHHPTGSRPTVGPPLHFGYS 60  
 DB 1 MKASGRVLVILICSVLSAVYIILCCWASLPCLATCLDHHPTGSRPTVGPPLHFGYS 60  
 QY 61 SVPDGKPLVREPCSCAVSSSGOMLGSGAGIIDSACVFRMNOAPTVGFADVGQRST 120  
 DB 61 SVPDGKPLVREPCSCAVSSSGOMLGSGAGIIDSACVFRMNOAPTVGFADVGQRST 120  
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 QY 181 VYFTEPRMAVCCQIFQDETGKRRSGSFLSTGTFMTLALBELCEIIVYGVMSDSYCR 240  
 DB 181 VYFTEPRMAVCCQIFQDETGKRRSGSFLSTGTFMTLALBELCEIIVYGVMSDSYCR 240  
 QY 241 EKSHPSVPHYFEKGRLEDCOMTLAHOAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300  
 DB 241 EKSHPSVPHYFEKGRLEDCOMTLAHOAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300

QY 301 TE 302  
 DB 301 TE 302

RESULT 3  
 CAG26699 PRELIMINARY; PRT; 302 AA.  
 AC CAG26699;  
 DT 23-APR-2004 (TrEMBLrel. 27, Created)  
 DT 23-APR-2004 (TrEMBLrel. 27, Last sequence update)  
 DE alpha-2,6-sialyltransferase ST6GalNac IV (Fragment).  
 GN ST6GALN4.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harduin-Leperc A., Martinez-Duncker I., Mollicone R., Delannoy P.,  
 RT "Phylogeny of sialyltransferases";  
 RL Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AJ646870; CAG26699.1; -;  
 KW Glycosyltransferase; transferase.  
 FT NON\_TER 1  
 FT NON\_TER 302 302  
 SQ SEQUENCE 302 AA; 34138 MW; AB7C984ABE53B553 CRC64;

Query Match 98.2%; Score 1597; DB 2; Length 302;  
 Best Local Similarity 98.0%; Pred. No. 5.5e-137;  
 Matches 296; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVPSAVYIILCCWAGLPCLATCLDHHPTGSRPTVGPPLHFGYS 60  
 DB 1 MKASGRVLVILICSVLSAVYIILCCWASLPCLATCLDHHPTGSRPTVGPPLHFGYS 60  
 QY 61 SVPDGKPLVREPCSCAVSSSGOMLGSGAGIIDSACVFRMNOAPTVGFADVGQRST 120  
 DB 61 SVPDGKPLVREPCSCAVSSSGOMLGSGAGIIDSACVFRMNOAPTVGFADVGQRST 120

QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 QY 181 VYFTEPRMAVCCQIFQDETGKRRSGSFLSTGTFMTLALBELCEIIVYGVMSDSYCR 240  
 DB 181 VYFTEPRMAVCCQIFQDETGKRRSGSFLSTGTFMTLALBELCEIIVYGVMSDSYCR 240  
 QY 241 EKSHPSVPHYFEKGRLEDCOMTLAHOAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300  
 DB 241 EKSHPSVPHYFEKGRLEDCOMTLAHOAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300  
 QY 301 TE 302  
 DB 301 TE 302

RESULT 4  
 ID Q704X3 PRELIMINARY; PRT; 302 AA.  
 AC Q704X3;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).  
 GN Name=slat7b; Synonyms=ST6GALNAC-IV;  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harduin-Leperc A., Martinez-Duncker I., Mollicone R., Delannoy P.,  
 RT "Phylogeny of sialyltransferases";  
 RL Submitted (Jan-2004) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AJ620652; CAF05851.1; -;  
 DR GO: 0047290; F: (alpha-N-acetylneuraminyl)-2,3-beta-galactosyl. . .; IEA.  
 DR GO: 0016757; F: transferase activity, transferring glycosyl. . .; IEA.  
 DR InterPro: IPR001675; Glyco\_transf\_29; 1.  
 KW Glycosyltransferase; transferase.  
 FT NON\_TER 1  
 FT NON\_TER 302 302  
 SQ SEQUENCE 302 AA; 34226 MW; 6C48E3977DA96733 CRC64;

Query Match 91.0%; Score 1479; DB 2; Length 302;  
 Best Local Similarity 90.1%; Pred. No. 3.1e-126;  
 Matches 272; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVPSAVYIILCCWAGLPCLATCLDHHPTGSRPTVGPPLHFGYS 60  
 DB 1 MKAPGRVLVILICSVLSAVYIILCCWASLPCLATCLDHHPTGSRPTVGPPLHFGYS 60  
 QY 61 SVPDGKPLVREPCSCAVSSSGOMLGSGAGIIDSACVFRMNOAPTVGFADVGQRST 120  
 DB 61 SVPDGKPLVREPCSCAVSSSGOMLGSGAGIIDSACVFRMNOAPTVGFADVGQRST 120  
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 QY 181 VYFTEPRMAVCCQIFQDETGKRRSGSFLSTGTFMTLALBELCEIIVYGVMSDSYCR 240  
 DB 181 VYFTEPRMAVCCQIFQDETGKRRSGSFLSTGTFMTLALBELCEIIVYGVMSDSYCR 240  
 QY 241 EKSHPSVPHYFEKGRLEDCOMTLAHOAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300  
 DB 241 EKSHPSVPHYFEKGRLEDCOMTLAHOAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300  
 QY 301 TE 302  
 DB 301 TO 302

```

RESULT 5
ID CAF05851 PRELIMINARY; PRT; 302 AA.
AC CAF05851;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).
GN ST6GALNAC-IV OR SIAT7B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Hardin-Leipers A., Martinez-Duncker I., Mollicone R., Delanny P.,
RA Oriol R.;
RT "Phylogeny of sialyltransferases."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620652; CAF05851.1; -.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 302 AA; 3426 MM; 6C48E3977DA96733 CRC64;

Query Match 91.8%; Score 1479; DB 2; Length 302;
Best Local Similarity 90.1%; Pred. No. 3.1e-126;
Matches 272; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHSGYS 60
DB 1 MKPGRLLIILICSVGFSAYIILCCWACLPFCPLATCLDPLSLINSRPTVPGPLHSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

RESULT 6
ID Q704X4 PRELIMINARY; PRT; 323 AA.
AC Q704X4;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).
GN Name=siat7b; Synonyms=st6galnac-iv;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hardin-Leipers A., Martinez-Duncker I., Mollicone R., Delanny P.,
RA Oriol R.;
RT "Phylogeny of sialyltransferases."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620651; CAF05850.1; -.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 323 AA; 36765 MM; 4BBF007BF6AAE481 CRC64;

Query Match 90.8%; Score 1476; DB 2; Length 323;
Best Local Similarity 91.1%; Pred. No. 6.3e-126;
Matches 275; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHSGYS 60
DB 1 MKPGRLLIILICSVGFSAYIILCCWACLPFCPLATCLDPLSLINSRPTVPGPLHSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

```

```

DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR001675; Glyco trans 29.
DR Pfam; PF00777; Glyco_transf_29; I.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 323 AA; 36765 MM; 4BBF007BF6AAE481 CRC64;

Query Match 90.8%; Score 1476; DB 2; Length 323;
Best Local Similarity 91.1%; Pred. No. 6.3e-126;
Matches 275; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHSGYS 60
DB 1 MKPGRLLIILICSVGFSAYIILCCWACLPFCPLATCLDPLSLINSRPTVPGPLHSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

RESULT 7
ID CAF05850 PRELIMINARY; PRT; 323 AA.
AC CAF05850;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).
GN SIAT7B OR ST6GALNAC-IV.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hardin-Leipers A., Martinez-Duncker I., Mollicone R., Delanny P.,
RA Oriol R.;
RT "Phylogeny of sialyltransferases."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620651; CAF05850.1; -.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 323 AA; 36765 MM; 4BBF007BF6AAE481 CRC64;

Query Match 90.8%; Score 1476; DB 2; Length 323;
Best Local Similarity 91.1%; Pred. No. 6.3e-126;
Matches 275; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHSGYS 60
DB 1 MKPGRLLIILICSVGFSAYIILCCWACLPFCPLATCLDPLSLINSRPTVPGPLHSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYPTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKRPPIVAFHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

```

```

Db      121 LAVIHTSVPLLRNNSHYFQOARDTLVYWGCGKMDRALGRTYRALLQLTRMVGQL 180
Qy      181 VYTFERRMAYACDQIFQDSTGKNRRSSGSLSTGWFMTLIALCEIIVYGVNSYCR 240
Db      181 VYTFERRMAYACDQIFQDSTGKNRRSSGSLSTGWFMTLIALCEIIVYGVNSYCR 240
Qy      241 EKSHSVPHVPEYKEKGLDECOMYLAEOAPRSARHFTTEKAVSPRAKRPVFAHPSMR 3000
Db      241 EKHHSVPHVPEYKEKGLDECOMYLAERAPRSARHFTTEKAVSPRAKRPVFAHPSMR 3000
Qy      301 TE 302
Db      301 TQ 302

RESULT 8
08C3J2
ID      08C3J2      PRELIMINARY;      PRT;      302 AA.
AC      08C3J2
DT      01-MAR-2003 (TRENBLrel. 23, Created)
DT      01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT      01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE      Mus musculus 10 days lactation, adult female mammary gland cDNA, RIKEN
DE      full-length enriched library, clone: D73003D08
DE      product: alpha1y1transferrase 7 (alpha-N-acetylnneuraminyl 2,3-
DE      beta-galactosyl-1,3)-N-acetyl galactosaminide alpha-2,6-
DE      beta1y1transferrase D, full insert sequence.
GN      Name=Slact7d;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX      NCBI_TaxId=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX      MEDLINE=99379253; PubMed=10349636;
RA      Carninci P., Hayashizaki Y.;
RT      "High-efficiency full-length cDNA cloning.";
RL      Meth. Enzymol. 303:19-44(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX      MEDLINE=21085660; PubMed=11217851;
RA      RIKEN FANTOM Consortium;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA      The FANTOM Consortium;
RT      "The FANTOM Genome Consortium Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs."
RL      Nature 420:563-573(2002).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX      MEDLINE=20499374; PubMed=11042159;
RA      Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA      Komoto H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT      "Normalization and subtraction of cap-trapper-selected cDNAs to
RT      prepare full-length cDNA libraries for rapid discovery of new genes.";
RL      Genome Res. 10:1617-1630(2000).
RN      [5]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX      MEDLINE=20530913; PubMed=11076861;
RA      Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA      Komoto H., Akiyama J., Nishi K., Kitesuna T., Tashiro H., Itoh M.,
RA      Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA      Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA      Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA      Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,

```



RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywiński M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC056451; AAHS6451.1; -  
 DR InterPro; IPR001675; Glyco trans\_29.  
 DR Pfam; PF00777; Glyco transf\_29; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 317 AA; 35740 MW; 61B8713C0234360D CRC64;

Query Match 89.3%; Score 1452; DB 2; Length 317;  
 Best Local Similarity 89.1%; Pred. No. 9.5e-124;  
 Matches 269; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 MKAPGRLVLTLLCSVPSAVYIILCCWAGLPCLATCLDHHFPGSRPTVPGRLHFGYS 60  
 DB 16 MKAPGRLVLTLLCSVPSAVYIILCCWAGLPCLATCLDHHFPAARSTVPGRLHFGYS 75  
 QY 61 SVPDGKPLVREPSCAVVSSSGOMLGSGLGAEIDSAECVFRNNOAPTYGFADVQORST 120  
 DB 76 SVPDGKPLVREPSCAVVSSSGOMLGSGLGAEIDSAECVFRNNOAPTYGFADVQORST 135  
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYVWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 136 LRVVSHTSVPLLRNYSHYFOKARDTLVYVWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 195  
 QY 181 VYFTFRMAVYCOIQIDETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYR 240  
 DB 196 VYFTFRMAVYCOIQIDETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYR 255  
 QY 241 EKSHPSVYHYFEKGRGLDECOMYLAEHOAPRSARHRTTEKAVFSRWAKKRPVFAHPSWR 300  
 DB 256 EKSPRSVYHYFEKGRGLDECOMYLAEHOAPRSARHRTTEKAVFSRWAKKRPVFAHPSWR 315  
 QY 301 TE 302  
 DB 316 AK 317

RESULT 10  
 AAHS6451 PRELIMINARY; PRT; 317 AA.  
 ID AAHS6451;  
 AC AAHS6451;  
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Slac7d protein (Fragment).  
 GN Slac7d.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywiński M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC056451; AAHS6451.1; -  
 DR InterPro; IPR001675; Glyco trans\_29.  
 DR Pfam; PF00777; Glyco transf\_29; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 317 AA; 35740 MW; 61B8713C0234360D CRC64;

Query Match 89.3%; Score 1452; DB 2; Length 317;  
 Best Local Similarity 89.1%; Pred. No. 9.5e-124;  
 Matches 269; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 MKAPGRLVLTLLCSVPSAVYIILCCWAGLPCLATCLDHHFPGSRPTVPGRLHFGYS 60  
 DB 16 MKAPGRLVLTLLCSVPSAVYIILCCWAGLPCLATCLDHHFPAARSTVPGRLHFGYS 75  
 QY 61 SVPDGKPLVREPSCAVVSSSGOMLGSGLGAEIDSAECVFRNNOAPTYGFADVQORST 120  
 DB 76 SVPDGKPLVREPSCAVVSSSGOMLGSGLGAEIDSAECVFRNNOAPTYGFADVQORST 135  
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYVWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
 DB 136 LRVVSHTSVPLLRNYSHYFOKARDTLVYVWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 195  
 QY 181 VYFTFRMAVYCOIQIDETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYR 240  
 DB 196 VYFTFRMAVYCOIQIDETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYR 255  
 QY 241 EKSHPSVYHYFEKGRGLDECOMYLAEHOAPRSARHRTTEKAVFSRWAKKRPVFAHPSWR 300  
 DB 256 EKSPRSVYHYFEKGRGLDECOMYLAEHOAPRSARHRTTEKAVFSRWAKKRPVFAHPSWR 315  
 QY 301 TE 302  
 DB 316 AK 317

RESULT 11  
 S17D MOUSE STANDARD; PRT; 360 AA.  
 ID S17D MOUSE  
 AC Q9R2B6; O88725; Q9JHP0; Q9QUP9; Q9R2B5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 05-JUL-2004 (Rel. 44, Last annotation update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3-N-acetyl-



Matches 266; Conservative 11; Mismatches 25; Indels 0; Gaps 0;

QY 1 MKRGRLLILITLCSVSAVYLLICWAGLPLCLATLDDHHPGSRPTVGPPLHFGYS 60  
 DB 1 MKRGRLLILITLCSVSAVYLLICWAGLPLCLATLDDHHPGSRPTVGPPLHFGYS 60

QY 61 SVDPGKRLTELCSCAVSSSGMGLGAGIDGAECLVLRNQAPTVGFEDVGRST 120  
 DB 61 SVDPGKRLTELCSCAVSSSGMGLGAGIDGAECLVLRNQAPTVGFEDVGRST 120

QY 121 LRVVSHTSVPLLRNYSHYQKARDITLYVWGGGRHMDRVLGRTYTTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYQKARDITLYVWGGGRHMDRVLGRTYTTLLQLTRMYPGLQ 180

QY 181 VYFTEERMAVYCDQIFODETKNRKRGSGSLSTGMFTMLALCEBIVVYGWVSDSYCR 240  
 DB 181 VYFTEERMAVYCDQIFODETKNRKRGSGSLSTGMFTMLALCEBIVVYGWVSDSYCR 240

QY 241 EKSHPSVYHYFEKGRLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPVFAHPSMR 300  
 DB 241 EKSHPSVYHYFEKGRLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPVFAHPSMR 300

QY 301 TE 302  
 DB 301 AK 302

RESULT 13

ID CAG26700 PRELIMINARY; PRT; 302 AA.

AC CAG26700;

DT 23-APR-2004 (TREMBLrel. 27, Created)

DT 23-APR-2004 (TREMBLrel. 27, Last sequence update)

DT 23-APR-2004 (TREMBLrel. 27, Last annotation update)

DE Alpha-2,6-sialyltransferase ST6GalNAc IV (Fragment).

GN ST6ATD.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Harduin-Lepers A., Martinez-Duncker I., Mollicone R., Delannoy P., Oriol R.

RT "Phylogeny of sialyltransferases";

RU Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; A0646871; CAG26700.1; -.

KW Glycosyltransferase; Transferase.

FT NON TER 1

FT NON TER 302

SO SEQUENCE 302 AA; 34332 MW; FA3F659B0C25886 CRC64;

Query Match 88.3%; Score 1436; DB 2; Length 302;  
 Best Local Similarity 88.1%; Pred. No. 2.6e-122;  
 Matches 266; Conservative 11; Mismatches 25; Indels 0; Gaps 0;

QY 1 MKRGRLLILITLCSVSAVYLLICWAGLPLCLATLDDHHPGSRPTVGPPLHFGYS 60  
 DB 1 MKRGRLLILITLCSVSAVYLLICWAGLPLCLATLDDHHPGSRPTVGPPLHFGYS 60

QY 61 SVDPGKRLTELCSCAVSSSGMGLGAGIDGAECLVLRNQAPTVGFEDVGRST 120  
 DB 61 SVDPGKRLTELCSCAVSSSGMGLGAGIDGAECLVLRNQAPTVGFEDVGRST 120

QY 121 LRVVSHTSVPLLRNYSHYQKARDITLYVWGGGRHMDRVLGRTYTTLLQLTRMYPGLQ 180  
 DB 121 LRVVSHTSVPLLRNYSHYQKARDITLYVWGGGRHMDRVLGRTYTTLLQLTRMYPGLQ 180

QY 181 VYFTEERMAVYCDQIFODETKNRKRGSGSLSTGMFTMLALCEBIVVYGWVSDSYCR 240  
 DB 181 VYFTEERMAVYCDQIFODETKNRKRGSGSLSTGMFTMLALCEBIVVYGWVSDSYCR 240

QY 241 EKSHPSVYHYFEKGRLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPVFAHPSMR 300  
 DB 241 EKSHPSVYHYFEKGRLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPVFAHPSMR 300

QY 301 TE 302  
 DB 301 AK 302

RESULT 14

ID Q704X2 PRELIMINARY; PRT; 289 AA.

AC Q704X2;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).

GN Name=slat7b; Synonyms=ST6GALNAc-IV;

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OX Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RA Harduin-Lepers A., Martinez-Duncker I., Mollicone R., Delannoy P., Oriol R.

RT Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; A0620653; CAF05852.1; -.

DR GO; GO:0047290; F:alpha-N-acetylneuraminyl-2,3-beta-galactos. . .; IEA.

DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.

DR InterPro; IPR001675; Glyco trans 29.

DR Pfam; PF00777; Glyco trans 29; I.

KW Glycosyltransferase; Transferase.

SO SEQUENCE 289 AA; 32780 MW; 9FB64C63BA56A74 CRC64;

Query Match 64.4%; Score 1047.5; DB 2; Length 289;  
 Best Local Similarity 66.3%; Pred. No. 6e-87;  
 Matches 195; Conservative 33; Mismatches 55; Indels 11; Gaps 2;

QY 6 RLVVILITLCSVSAVYLLICWAGLPLCLATLDDHHPGSRPTVGPPLHFGYSVDPG 65  
 DB 4 RLVVILITLCSVSAVYLLICWAGLPLCLATLDDHHPGSRPTVGPPLHFGYSVDPG 65

QY 66 KPLVREPCSCAVSSSGMGLGAGIDGAECLVLRNQAPTVGFEDVGRSTRVYS 125  
 DB 53 KPLVREPCSCAVSSSGMGLGAGIDGAECLVLRNQAPTVGFEDVGRSTRVYS 112

QY 126 HTSVPLLRNYSHYQKARDITLYVWGGGRHMDRVLGRTYTTLLQLTRMYPGLQVYFT 185  
 DB 113 HTSVPLLRNYSHYQKARDITLYVWGGGRHMDRVLGRTYTTLLQLTRMYPGLQVYFT 172

QY 186 ERMMAVYCDQIFODETKNRKRGSGSLSTGMFTMLALCEBIVVYGWVSDSYCRKSH 245  
 DB 173 ERMMAVYCDQIFODETKNRKRGSGSLSTGMFTMLALCEBIVVYGWVSDSYCRKSH 232

QY 246 SVFHYHFEKGRLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPVFAHPSMR 299  
 DB 233 SVFHYHFEKGRLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPVFAHPSMR 286

RESULT 15

ID CAF05852 PRELIMINARY; PRT; 289 AA.

AC CAF05852;

DT 02-MAR-2004 (TREMBLrel. 27, Created)

DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)

DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)

DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).

GN ST6GALNAc-IV OR ST6AT7B.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;



**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:05:03 ; Search time 40 Seconds  
(without alignments)  
726.437 Million cell updates/sec

Title: US-10-019-114A-7

Sequence: 1 MKRPGRLVLIICSVFSAV.....FSRMKKRPVFAHPSWRT 302

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR 79: \*  
2: PIR1: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	239	14.7	350	2	B54420
2	230	14.1	350	2	JC5251
3	228.5	14.1	350	2	A54420
4	221.5	13.6	524	2	JC8017
5	209.5	12.9	337	2	S36824
6	206.5	12.7	342	2	S55675
7	202.5	12.5	380	2	UC6321
8	200.5	12.3	359	2	JC4224
9	198.5	12.2	340	2	I54229
10	196.5	12.1	359	2	S52425
11	191	11.7	343	2	A45073
12	190.5	11.7	380	2	A56950
13	188.5	11.6	359	2	I59403
14	182.5	11.2	413	2	S41114
15	178	10.9	566	2	A49880
16	172	10.6	374	2	A45074
17	172	10.6	375	2	JN0618
18	168	10.3	406	2	A41734
19	164.5	10.1	356	2	A54032
20	164.5	10.1	526	2	UC7248
21	158	9.7	403	2	A28451
22	157	9.7	359	2	UC4973
23	155.5	9.6	342	2	JC5195
24	153.5	9.4	425	2	T00720
25	147	9.0	398	2	G86216
26	144.5	8.9	375	2	I48666
27	143	8.8	329	2	A48715
28	143	8.8	332	2	A49879
29	141.5	8.7	375	2	A46727

30	140.5	8.6	376	2	JC5600	alpha-N-acetylneur
31	139.5	8.6	375	2	I39169	siatyltransferase
32	137.5	8.5	404	2	A54871	Gal beta-1, 3GalNA
33	134	8.2	387	2	JB0364	lactosylceramide a
34	127.5	7.8	222	2	B54898	STX protein - huma
35	126	7.7	470	2	T49278	hypothetical prote
36	97	6.0	255	2	T52352	hypothetical prote
37	93.5	5.8	990	2	D83706	antibiotic mersac
38	89	5.5	431	2	B37802	ctx protein - Erw
39	88.5	5.4	883	2	A96662	hypothetical prote
40	88	5.4	480	2	S14394	bindin precursor -
41	86.5	5.3	517	2	S32169	hypothetical prote
42	86	5.3	618	2	I48914	protein-tyrosine k
43	85.5	5.3	461	2	B41313	probable transposa
44	84	5.2	322	2	S66766	probable membrane
45	84	5.2	642	1	S52111	uromodulin precurs

## ALIGNMENTS

```
RESULT 1
B54420
beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) ST3GalA.2 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C:Accession: B54420
R:Lee, Y.C.; Kojima, N.; Wada, E.; Kurosawa, N.; Nakacka, T.; Hamamoto, T.; Tsuji, S.
J. Biol. Chem. 269, 10028-10033, 1994
A:Title: Cloning and expression of cDNA for a new type of Galbeta1,6GalNAc alpha2,3-sialyl
A:Cross-references: UNIPROT:Q11205; GB:X76988; NID:9475225; PIDN:CAA54293.1; PID:9475226
A:Reference number: A54420; MUID:94193584; PMID:8144500
A:Accession: B54420
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-350 <LEE>
C:Superfamily: beta-galactoside alpha-2,3-sialyltransferase
C:Keywords: glycosyltransferase; transmembrane protein

Query Match 14.7%; Score 239; DB 2; Length 350;
Best Local Similarity 34.3%; Pred. No. 2,2e-13;
Matches 71; Conservative 24; Mismatches 80; Indels 32; Gaps 8;

Cy 59 YSSVPDCKP-VREP--CRCAVYSSGOMLGSLGAEDISAEVFMNQPTTGFADY 115
Db 132 FOYVGENPYRFRDPQCCRCVAVGNSGNLRGSGYGVDSHNFIMMNOAPVGFEDV 191
Cy 116 GORSTLRVSGHTSVPLLRVYSHYFQ--KARDLYMV-----WGQGRMDRVLGGRYRTL 169
Db 192 GSRTHHFMTPEBAKULPANVSFVLVFPKALDLMWIMASLSTQIR-----TAPV 243
Cy 170 LQTRMYPGLOVYTFERRMAVYCDQIFQDETGNRRQSGFLSTGFTMTLALCEIV 229
Db 244 KSFLRV-----DKEKQVLYNPAPFKYIHDRWTEHHGRVPSGMLVLFALHVCDEVN 295
Cy 230 VYGVNDSYCREKSHPSVPYHYFEKGR 256
Db 296 VYGFAGDS--RGWNR-----HYWENNR 315

RESULT 2
JC5251
beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - human
C:Species: Homo sapiens (man)
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: JC5251; G01021
R:Kim, Y.T.; Kim, K.S.; Kim, S.H.; Kim, C.H.; Ko, J.H.; Choe, I.S.; Tsuji, S.; Lee, Y.C.
Biochem. Biophys. Res. Commun. 228, 324-327, 1996
A:Title: Molecular cloning and expression of human Gal beta 1,3GalNAc alpha 2,3-sialyltra
A:Reference number: JC5251; MUID:97079181; PMID:8920913
A:Accession: JC5251
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
```

A;Residues: 1-350 <KIM>  
 A;Cross-references: UNIPROT:Q16842; GB:U63090; NID:G1773282; PIDN:AA040389.1; PID:G17732  
 A;Experimental source: liver  
 R;Jordanengo, V.  
 Submitted to the EMBL Data Library, March 1996  
 A;Reference number: H00561  
 A;Accession: G01021  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-350 <GIO>  
 A;Cross-references: EMBL:X96667; NID:G1235530  
 C;Comment: This enzyme catalyzes the transfer of sialic acid from CMP-NeuAc to the terminal  
 substrate preference for glycolipid than for O-linked oligosaccharides of glycoproteins.  
 C;Genetics:  
 A;Gene: ST3 (0) -11  
 C;Superfamily: beta-galactoside alpha-2,3-sialyltransferase  
 C;Keywords: glycosyltransferase

Query Match 14.1%; Score 230; DB 2; Length 350;  
 Best Local Similarity 33.3%; Pred. No. 1,3e-12;  
 Matches 69; Conservative 24; Mismatches 82; Indels 32; Gaps 8;

QY 59 YSSVDPGKPL-VREP--CRSCAVVSSGQMLGSLGAELIDSAECVFRMNOAPTVGFADY 115  
 DB 132 FOIVPGENPFRFPDPOCRCAVVGNSGNLRSGYGVGDVDGHNFMIMNOAPTVGFEDV 191  
 QY 116 GQSTLRVSHTSVPLLRVSHYFO--KARDTLVWV---WGGRHMDVLAGRTYRTL 169  
 DB 192 GSRTHHFMPESAKNLPANVSFVLVPFKVLDLWIMASLSTGQIR-----TYAPV 243  
 QY 170 LQITRMYPGLQVYTFTRMMAVCDQIFODETGKNNRSGSFLSTGWTMLALECEIV 229  
 DB 244 KSEFLRV-----DKKKQIYNPAFFKYIHDRWTEHNGRYPRSTGMVLVLFALHVCDEV 295  
 QY 230 VYGVSDSYCREKSHPSVPYHYEKGK 256  
 DB 296 VYGFAGDS--RGNWH-----HYWENNR 315

## RESULT 3

A54420  
 beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) ST3GAL.2 - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004  
 C;Accession: A54420  
 R;Lee, Y.C.; Kojima, N.; Mada, E.; Kurosawa, N.; Nakaoka, T.; Hamamoto, T.; Tsuji, S.  
 J. Biol. Chem. 269, 10028-10033, 1994  
 A;Title: Cloning and expression of cDNA for a new type of Galbeta1,6Galnac alpha2,3-sial  
 A;Reference number: A54420; MUID:94193584; PMID:8144500  
 A;Accession: A54420  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-350 <LEE>  
 A;Cross-references: UNIPROT:Q11204; GB:X76989; NID:G475203; PIDN:CAA54294.1; PID:G475204  
 C;Superfamily: beta-galactoside alpha-2,3-sialyltransferase  
 C;Keywords: glycosyltransferase; transmembrane protein

Query Match 14.1%; Score 228.5; DB 2; Length 350;  
 Best Local Similarity 34.6%; Pred. No. 1.8e-12;  
 Matches 65; Conservative 23; Mismatches 75; Indels 25; Gaps 6;

QY 59 YSSVDPGKPL-VREP--CRSCAVVSSGQMLGSLGAELIDSAECVFRMNOAPTVGFADY 115  
 DB 132 FOIVPGENPFRFPDPOCRCAVVGNSGNLRSGYGVGDVDGHNFMIMNOAPTVGFEDV 191  
 QY 116 GQSTLRVSHTSVPLLRVSHYFO--KARDTLVWV---WGGRHMDVLAGRTYRTL 169  
 DB 192 GSRTHHFMPESAKNLPANVSFVLVPFKVLDLWIMASLSTGQIR-----TYAPV 243  
 QY 170 LQITRMYPGLQVYTFTRMMAVCDQIFODETGKNNRSGSFLSTGWTMLALECEIV 229  
 DB 244 KSEFLRV-----DKKKQIYNPAFFKYIHDRWTEHNGRYPRSTGMVLVLFALHVCDEV 295

QY 230 VYGVSDS 237  
 DB 296 VYGFAGDS 303

## RESULT 4

beta-galactoside alpha-2,6-sialyltransferase (EC 2.4.99.1) II - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 04-Apr-2004 #sequence\_revision 04-Apr-2004 #text\_change 04-Apr-2004  
 C;Accession: JCB017  
 R;Takahama, S.; Tsuji, S.; Tsujimoto, M.  
 J. Biochem. 134, 287-296, 2003  
 A;Title: Comparison of the enzymatic properties of mouse beta-galactoside alpha2,6-sialy  
 A;Reference number: JCB017; PMID: 12966079  
 A;Accession: JCB017  
 A;Molecule type: mRNA  
 A;Residues: 1-524 <TKA>  
 A;Cross-references: DDBJ:AB095093  
 C;Comment: This enzyme is an oligosaccharide-specific enzyme and a type II transmembrane  
 cid to galactose of the Gal beta1,4 GlcNAc structure through an alpha 2,6-linkage.  
 C;Genetics:  
 A;Gene: ST6GALII  
 A;Map position: 17C  
 A;Intons: 310/1; 342/3; 376/3; 435/1  
 C;Keywords: oligosaccharide-specific; sialyl motif; sialyltransferase; ST6GALII; transmen

Query Match 13.6%; Score 221.5; DB 2; Length 524;  
 Best Local Similarity 28.6%; Pred. No. 1.2e-11;  
 Matches 64; Conservative 38; Mismatches 83; Indels 39; Gaps 8;

QY 45 GSRPTVPG-----PLHPSGVSVDPGKPLVREPGRSCAVVSSGQMLGSLGAELIDSAECV 100  
 DB 270 GMRPLVPGVPLSGLHRLGSL-----SCAVVMSAGAILNSSLGSEIDSHDAV 315  
 QY 101 FRMNOAPTVGFADVQGRSTLRVSHTSVPLLRVSHYFOQA--RDTLYWVGGRHM 157  
 DB 316 LRFNSAPRTGYEKVDGNKKTVRIINSQ---ILANSHHFIDSLAKYDVLVMDAPAY 371  
 QY 158 DRV-----LGGTYRTIQLQITRMYPGLQVYTFTRMMAVCDQIFODETGKNNRSGSFL 211  
 DB 372 ANNLNMYKKPDYNLFPFYIQRKRYPTQPIYLPFIQLMDIENREKIQPNP-- 429

QY 212 STGMFTMLALECEIVVYGVSDSYCREKSHPSVYH--YPE 253  
 DB 430 SSGFGLIMSMCKEVHYEYIPSVRQTELC-----YHELYVD 469

## RESULT 5

beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 22-Jan-1994 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C;Accession: S36824  
 R;Lee, Y.C.; Kurosawa, N.; Hamamoto, T.; Nakaoka, T.; Tsuji, S.  
 Eur. J. Biochem. 216, 377-385, 1993  
 A;Title: Molecular cloning and expression of Gal-beta1,3Galnac-alpha-2,3-sialyltransfer  
 A;Reference number: S36824; MUID:93387288; PMID:8375377  
 A;Accession: S36824  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-337 <LEE>  
 A;Cross-references: UNIPROT:P54751; EMBL:X73523; NID:G402214; PIDN:CAA51919.1; PID:G40221  
 C;Superfamily: beta-galactoside alpha-2,3-sialyltransferase  
 C;Keywords: glycosyltransferase

Query Match 12.9%; Score 209.5; DB 2; Length 337;  
 Best Local Similarity 32.6%; Pred. No. 8.3e-11;  
 Matches 61; Conservative 32; Mismatches 65; Indels 29; Gaps 7;

QY 73 CRSCAVVSSGQMLGSLGAELIDSAECVFRMNOAPTVGFADVQGRSTLRVSHTSVPL 132  
 DB 136 CRCAVVGNSGNLKDSSYGEIDSHDFVLRMNKAPTVGFADVGSRTTHLV----- 187



Query Match 12.7%; Score 206.5; DB 2; Length 342;  
 Best Local Similarity 28.1%; Pred. No. 1.5e-10;  
 Matches 61; Conservative 37; Mismatches 68; Indels 51; Gaps 7;

133 LNAVSHYFOKARDTLVWVGQGRHMDR--VLGGRTYRTLLQLTRMY-----PGLQVTFTE 186  
 188 ----YPSFRRLGNNVNVLPFKTDLQWISATTTCT---IHTYVPVPRKIKV--QE 239

187 RMAYCDQIFQDETGNKRRSGSFLSTGWTMLALCEIIVVYGMVSDSYCREKSHPS 246  
 240 KILYHPAFIKYVFDWMLQGHGRYPTGLIIFSIHCDEVDLYGFGADS--KGNW-- 295

247 VPHYFE 253  
 296 --HYME 299

## RESULT 6

Gal-beta-1,3galNac alpha-2,3-sialyltransferase - chicken  
 C/Species: Gallus gallus (chicken)  
 C/Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
 C/Accession: S55675

R;Kurosawa, N.; Hamamoto, T.; Inoue, M.; Tsuji, S.  
 Biochim. Biophys. Acta 1244, 216-222, 1995  
 A/Title: Molecular cloning and expression of chick Gal-beta-1,3galNac alpha-2,3-sialyltr  
 A/Reference number: S55675; PMID:95284088; PMID:7766661  
 A/Accession: S55675  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1342 <KUR>  
 A/Cross-references: UNIPROT:Q11200; GB:X80503; NID:g975654; PID:CAA5666.1; PID:g975655  
 C/Superfamily: beta-galactoside alpha-2,3-sialyltransferase

Query Match 12.7%; Score 206.5; DB 2; Length 342;  
 Best Local Similarity 28.1%; Pred. No. 1.5e-10;  
 Matches 61; Conservative 37; Mismatches 68; Indels 51; Gaps 7;

50 VPG---PLHPSGYSPDGPPLVREPCRSQAVSSSGOMLGSGLAEIDSACVFRNQA 106  
 126 IEDDROPLQGRGTF-----CRCAVVGNSGNLRSQYQGDIDSHDFLAWNRA 174

107 PTMGFEADVQGRSTLRVSHTSVPLLRNYSHTFQKARDTLVWVGQGRHMDRYLGRIT 166  
 175 PTYGESDVGSXTTHFV-----YPSYKEIAENVSMI-----VIPFKTL 214

167 RTLLQLTRMYPGIQLVTF-----ERMAYCDQIFQDETGNKRRSGSFLSTGWF 216  
 215 DLRWIVTALTTGTFVTPVPRKIKVREKVKLVLPNPSFKTYVENMLQNHGRYPTGL 274

217 TMLALELCEIIVVYGMVSDSYCREKSHPSVPHYFE 253  
 275 SVTFALHVCDEVNVYGRGADS---KGNW---WHHYME 304

## RESULT 7

JC6321  
 CMP-NeuAc-GM3 alpha2-8 sialyltransferase (EC 2.4.99.-) [imported] - rat  
 N/Alternate names: GM3-synthase; GM3-synthase  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 14-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
 C/Accession: JC6321  
 R;Zeng, G.; Gao, L.; Yu, R.K.  
 Gene 187, 131-134, 1997  
 A/Title: Cloning of the cDNA coding for rat brain CMP-NeuAc:GM3 alpha2-8 sialyltransfera  
 A/Reference number: JC6321; PMID:97225806; PMID:9073076  
 A/Accession: JC6321  
 A/Molecule type: mRNA  
 A/Residues: 1-380 <ZEN>  
 A/Cross-references: UNIPROT:P97877; GB:U55938  
 A/Experimental source: brain  
 C/Comment: This enzyme is one of the sialyltransferases and adds one more alpha-2,8-link  
 C/Superfamily: alpha-N-acetylnneuraminate alpha-2,8-sialyltransferase  
 C/Keywords: glycosyltransferase  
 F;18-33/Domain: transmembrane #status predicted <TM>

Query Match 12.5%; Score 202.5; DB 2; Length 380;  
 Best Local Similarity 30.4%; Pred. No. 4e-10;  
 Matches 66; Conservative 29; Mismatches 79; Indels 43; Gaps 8;

60 SSVPDQKPLVREPCRSQAVSSSGOMLGSGLAEIDSACVFRNQAFTVGFADVGRS 119  
 146 SLIPDSVPIIKRKNVCAVVGNSGILTGSCGGEIDKSPVFCNPAFTAFAHDKYKT 205

120 TLRVSHTSVPLLRNYSHTFQKARDTL-----YMWGQGRHMDRVLGRIT 168  
 206 NLT---TENPSILEKTYNNLLTIQDRNFFLSKLKDLGAILW-----IPAFFHT 252

169 ILQITMYPGIQLVTFE-----MMAYCDQIFQ--DETGNKRRSGSFLSTGWTML 220  
 253 SATVTRTL-----VDFVEHGRGLKQLAMPGNIMGVNRYWKOKHLSPKRLSTGFLVYTL 308

221 ALELCBIIVVYGMV---DSYCREKSHPSVPHYFEK 254  
 309 ASALCEIHLYGFWPFGFDPNTR---DLPHYDYK 341

## RESULT 8

JC4224  
 alpha-N-acetylnneuraminate alpha-2,8-sialyltransferase (EC 2.4.99.8) - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 19-Oct-1995 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004  
 C/Accession: JC4224  
 R;Yoshida, Y.; Kojima, N.; Tsuji, S.  
 J. Biochem. 118, 658-664, 1995  
 A/Title: Molecular cloning and characterization of a third type of N-glycan alpha 2,8-si  
 A/Reference number: JC4224; PMID:96115941; PMID:8690732  
 A/Accession: JC4224  
 A/Molecule type: mRNA  
 A/Residues: 1359 <YOS>  
 A/Cross-references: UNIPROT:Q64692; EMBL:X86000; NID:g2665331; PID:CAA59992.1; PID:g122;  
 A/Experimental source: lung  
 C/Comment: This enzyme exhibits activity specific toward sialylated glycoproteins. This e  
 C/Superfamily: alpha-N-acetylnneuraminate alpha-2,8-sialyltransferase  
 C/Keywords: glycoprotein; glycosyltransferase; lung; transmembrane protein  
 F;141-185/Region: sialyl signature L  
 F;280-302/Region: sialyl signature S  
 F;50,74,119,204,219/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match 12.3%; Score 200.5; DB 2; Length 359;  
 Best Local Similarity 27.9%; Pred. No. 5.5e-10;  
 Matches 70; Conservative 38; Mismatches 98; Indels 45; Gaps 9;

59 YSVPDQKPLVREPCRSQAVSSSGOMLGSGLAEIDSACVFRNQAFTVGFADVGR 118  
 125 HSLLPEVSPKRNRRFTCAVVGNSGILIDSGCKEIDSHNFVRLCLAPVVERADVGT 184

119 SLRVVSHTSVPLLRNYS-----HYFOKARDTL-----YMWGQGRHMDRVLGR 164  
 185 SDPTVNPVSVQAFAGFGRNEDREKFRHLSTLNDVSLVMI PAFVYVYGEKHEVW--NAL 243

165 TYRTLLQLTRMYPGIQLVTFE-----YTERMMAYCDQIFQDETGNKRRSGSFLSTGFTM 218  
 244 ILNKKLQVTRVPSLRLIHAVRGYWLTNKKVP-----KRP-----STGLMY 265

219 ILALELCEIIVVYGMVSDSYCREKSHPSVPHYFEKGRLDCEQMYLAHQAPSAARFIT 278  
 286 TLMTRCDEHLHYGFW--PPKDLNKAAYKTHYD---DLKTRYFSNAPSHPMPLEFKT 339

279 EKAVFSRWAKK 289  
 340 LNVTLNHRGALK 350

## RESULT 9

beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - human  
 C/Species: Homo sapiens (man)



[illegible]

RESULT 13  
159403  
alpha-2,8-poly(sialyl)transferase - human  
C:Species: Homo sapiens (man)  
C:Date: 31-May-1996 #sequence  
C:Accession: 159403  
R:Nakayama, J.; Fukuda, M.N.; Fredette, B.; Ranscht, B.; Fukuda, M.  
Proc. Natl. Acad. Sci. U.S.A. 92, 7031-7035, 1995  
A:Title: Expression cloning of a human poly(sialyl)transferase that forms the poly(sialyl)  
A:Reference number: 159403; MUID:95350205; PMID:7623364  
A:Accession: 159403  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: mRNA  
A:Residues: 1-359 <RSC>  
A:Cross-references: UNIPROT:Q92187; GB:I41680; NID:G945220; PIDN:AAC41775.1; PID:G945221  
C:Genetics:  
A:Gene: GDB:IPST  
A:Cross-references: GDB:I327254  
C:Superfamily: alpha-N-acetylneuraminatase alpha-2,8-sialyltransferase

[illegible]

RESULT 14  
S4114  
Gal beta 1,4 GlcNAc alpha 2,6-sialyltransferase - chicken  
C1:Species: Gallus gallus (chicken)  
C1:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C1:Accession: S4114  
E1:Kuroiwa, N.; Kawasaki, M.; Hamamoto, T.; Nakaoaka, T.; Lee, Y.-C.; Arita, M.; Teuji, S

Bur. J. Biochem. 219, 375-381, 1994  
 A>Title: Molecular cloning and expression of chick embryo Gal-beta-1,4GlcNAc-alpha-2,6-sialyltransferase  
 A.Reference number: S81114; MUID:94139712; PMID:8507003  
 A.Accession: S81114  
 A>Status: Preliminary  
 A.Molecule type: mRNA  
 A.Residues: 1-413 <KUR>  
 A.Cross-references: UNIPROT:Q92182, EMBL:X75558, NID:q452750, PIDD:q45275.1, PIDD:q45275.1, PIDD:q45275.1  
 A.Superfamily: beta-galactoside alpha-2,6-sialyltransferase

	Query Match	Similarity	Score	DB 2	length
Best Local	61	26.1%	182.5	DB 2	413
Matches	61	Conservative	39	Mismatches	95
				Indels	39
				Gaps	8

RESULT 15  
A49880  
alpha-N-acetylglactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.3) - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 30-Jun-1995 #sequence\_revision 30-Jun-1995 #ext\_change 09-Jul-2004  
C:Accession: A49880  
R:Kurosawa, N.; Hamamoto, T.; Lee, Y.C.; Nakaoka, T.; Kojima, N.; Tsuji, S.  
J. Biol. Chem. 269, 1402-1409, 1994  
A:Title: Molecular cloning and expression of GalNAc alpha2,6-sialyltransferase.  
A:Reference number: A49880; PMID:94117458; PMID:8288607  
A:Accession: A49880  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-566 <KUR>  
A:Cross-references: UNIPROT:Q92183; GB:X74946; NID:G453196; PIDD:CA52902.1; PIDD:G453197  
C:Superfamily: alpha-N-acetylglactosaminide alpha-2,6-sialyltransferase  
C:Keywords: glycosyltransferase

	Query March	10.9%;	Score 178;	DB 2;	Length 566;	
	Best Local Similarity	25.2%;	Pred. No. 9,3e-08;			
	Matches	63;	Conservative	37;	Mismatches	80;
					Indels	70;
					Gaps	9;
Qy	73	CRSCAVSSSSCOMLGGSLGAEDISACVPRMNOAPTGVGEADVGORSTLRVSHSVPLT	132			
Dd	337	CISCAGVANGGILNNSGMGOEIDSHDYFRVSGAVIKGYEKDVGTSFPGFAAYSLVSS	396			
Qy	133	LRYNTSHY-FOKARDTLVWVGGRHMDRVLGSRTRYRTLLQTETMYPCGLQV-----	181			
Dd	397	LQNLGHKRFKKI-----PQGKH-----RTHPLEAVADYEWLKALLLDKOIRKG	441			
Qy	182	-----YPTFERMAVYCDOIIFODETGKAFRRQGSFL-----STGW	215			
Dd	442	FLANYGRPRRERFDEFTMNKKLYLAHPDL---RYLKNNFLFKSKNQKPYRWRLYRPPTGA	498			
Qy	216	FTMLIALCEBEIVVYGWSDSYCREKSHPSVPYHFEKGRLDCCQMYLLAHQAAPRAHR	275			
Dd	499	LLLTALHLCDRVSAVGYITE-----GHQKYSDDHYDK-EWRKLVEFYVNH-----	543			
Qy	276	FTEKAVERFSR	285			
Dd	544	FNLEKQVMKR	553			

Tue Dec 14 09:16:19 2004

us-10-019-114a-7.rpr

Page 6

Search completed: December 10, 2004, 18:25:57  
Job time : 41 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: December 21, 2004, 18:02:40 ; Search time 25 Seconds

(without alignments)  
801.121 Million cell updates/sec

Title: US-10-019-114A-7

Perfect score: 1626

Sequence: 1 MKAPGRLVLIILICSVFSAV.....FSRWAKKRPVPAHPSWRB 302

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*
- 4: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*
- 5: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*
- 6: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1626	100.0	302 4 US-09-599-360B-75	Sequence 75, App1
2	639.5	39.3	336 3 US-09-334-601-13	Sequence 13, App1
3	494	30.4	95 4 US-09-621-976-3915	Sequence 3915, App1
4	198.5	12.2	340 2 US-08-446-875-16	Sequence 16, App1
5	196.5	12.1	359 1 US-08-503-133A-2	Sequence 2, App1
6	196.5	12.1	359 2 US-08-576-775A-2	Sequence 2, App1
7	196.5	12.1	359 3 US-08-972-448-2	Sequence 2, App1
8	196.5	12.1	359 4 US-08-999-445-2	Sequence 2, App1
9	191	11.7	343 2 US-08-446-875-2	Sequence 2, App1
10	191	11.7	343 3 US-08-102-385G-2	Sequence 2, App1
11	190.5	11.7	339 1 US-08-626-994A-3	Sequence 3, App1
12	190.5	11.7	339 2 US-08-957-742-3	Sequence 3, App1
13	190.5	11.7	364 1 US-08-626-994A-1	Sequence 1, App1
14	190.5	11.7	364 2 US-08-957-742-1	Sequence 1, App1
15	178	10.9	355 2 US-08-666-367B-6	Sequence 6, App1
16	178	10.9	355 3 US-09-143-438-6	Sequence 6, App1
17	178	10.9	566 2 US-08-666-367B-5	Sequence 5, App1
18	178	10.9	566 3 US-09-143-438-5	Sequence 5, App1
19	174.5	10.7	329 3 US-09-334-601-10	Sequence 10, App1
20	173.5	10.7	90 3 US-09-334-601-21	Sequence 21, App1
21	172	10.6	375 2 US-08-446-875-10	Sequence 10, App1
22	172	10.6	375 3 US-08-102-385G-10	Sequence 10, App1
23	171	10.5	376 2 US-08-666-367B-8	Sequence 8, App1
24	171	10.5	376 3 US-09-143-438-8	Sequence 8, App1
25	170	10.5	374 2 US-08-446-875-4	Sequence 4, App1
26	170	10.5	374 3 US-08-102-385G-4	Sequence 4, App1
27	168	10.3	90 3 US-09-334-601-20	Sequence 20, App1

28	168	10.3	406 1 US-08-446-777-4	Sequence 4, App1
29	168	10.3	767 1 US-08-446-777-6	Sequence 6, App1
30	168	10.3	30 3 US-08-446-777-8	Sequence 8, App1
31	166.5	10.2	90 3 US-09-334-601-22	Sequence 22, App1
32	166.5	10.2	90 3 US-09-334-601-24	Sequence 24, App1
33	165	10.1	600 4 US-10-140-002-462	Sequence 462, App1
34	163.5	10.1	356 4 US-08-361-304A-2	Sequence 2, App1
35	159	9.8	340 2 US-08-102-385G-18	Sequence 18, App1
36	157.5	9.7	90 3 US-09-334-601-23	Sequence 23, App1
37	148.5	9.1	90 3 US-09-334-601-25	Sequence 25, App1
38	145.5	8.9	329 1 US-07-991-587A-7	Sequence 26, App1
39	143	8.8	329 1 US-08-309-985-7	Sequence 7, App1
40	143	8.8	332 2 US-08-446-875-12	Sequence 12, App1
41	143	8.8	332 2 US-08-102-385G-12	Sequence 12, App1
42	143	8.8	384 3 US-09-334-601-8	Sequence 8, App1
43	142.5	8.8	77 2 US-08-102-385G-29	Sequence 29, App1
44	142.5	8.8	55 2 US-08-102-385G-15	Sequence 15, App1
45	142	8.7		

## ALIGNMENTS

```
RESULT 1
US-09-599-360B-75
; Sequence 75, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bouguerelet, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 75
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -18..-1
; US-09-599-360B-75

Query Match      100.0% Score 1626; DB 4; Length 302;
Best Local Similarity 100.0%; Pred No. 3.4e-170;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKAPGRLVLIILICSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGELHSGYS 60
DB      1 MKAPGRLVLIILICSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGELHSGYS 60
QY      61 SVPDGRLVREPRSCAVVSSSGOMIGSGIAGETDABECVFRNNOAPTYGFEDVQGRST 120
DB      61 SVPDGRLVREPRSCAVVSSSGOMIGSGIAGETDABECVFRNNOAPTYGFEDVQGRST 120
QY      121 LRVVSHTSVPLLRNYSYHFOKARDTLVYVWGGGRHMDVLAGRTYRTTLQLTRMYPGLQ 180
DB      121 LRVVSHTSVPLLRNYSYHFOKARDTLVYVWGGGRHMDVLAGRTYRTTLQLTRMYPGLQ 180
QY      181 VYTFERRMAVCDQIFQDETGKRRQSGFPLSTGFTMTLALCEIIVYVGVSDSYCR 240
DB      181 VYTFERRMAVCDQIFQDETGKRRQSGFPLSTGFTMTLALCEIIVYVGVSDSYCR 240
QY      241 EKSHPSVPPIHYEKGKLDCEQNTLAHBOAPRSHRITTEKAVPSRAKXRPVFAHPSWR 300
DB      241 EKSHPSVPPIHYEKGKLDCEQNTLAHBOAPRSHRITTEKAVPSRAKXRPVFAHPSWR 300
```

Db 241 EKSHPSVHYFPEKGLDECOMYLAHQAPRSARHPTTEKAVSRWAKRPVFAHPSWR 300  
QY 301 TE 302  
Db 301 TE 302

## RESULT 2

US-09-334-601-13  
Sequence 13, Application US/09334601  
Patent No. 6280989  
GENERAL INFORMATION:  
APPLICANT: Kapitonov, Dmitri  
APPLICANT: Yu, Robert  
TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES  
FILE REFERENCE: VCUIP-6  
CURRENT APPLICATION NUMBER: US/09/334,601  
CURRENT FILING DATE: 1999-06-17  
NUMBER OF SEQ ID NOS: 94  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 336  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-334-601-13

Query Match 39.3%; Score 639.5; DB 3; Length 336;  
Best Local Similarity 47.9%; Pred. No. 1,1e-61;  
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPVTP-GEHFSGYSSVPDGRPLVREPCRSQAVVSSSGOMLGSLGAEIDSAECYFR 102  
Db 64 TQGRPGVAGRPRLDGYLVADHKPL-KMCRDCAVTSSGHLHSRQSGQIDQTECVIR 122  
QY 103 MNQAPVGFENDVQGRSTLRVSHTSVPLLRNYSHFQKARDTLVWVGGRMDRVLG 162  
Db 123 MNDPRTGYGRDVQNRSLRYLHSSIORILRNHDLINSSQGVTFIFWGPSSYMRDGR 182  
QY 163 GRVYRTLLQLTRMYPGLOVYTFTERMAVCDQIFODETGKRRQSGSFLSTGWTMLAL 222  
Db 183 GQVYNNHLHLSQVLRKAKFWITRHKMLQFDELFKQETGKORIKISNTWLSTGWTMTIAL 242  
QY 223 ELCEEIVVYGWSSGYCRKSHSPVRYHYFEKGLDECOMYLAHQAPR-SARHPTTEKA 281  
Db 243 ELCDRIIVYGWVPPDPFCRDPNHPSPVRYHYEPGPRDECTMYLSHERGKRGSHRPTTEKR 302  
QY 282 VFSRWAKRPVFAHPSWRTE 302  
Db 303 VFSRWAKRPVFAHPSWRTE 323

## RESULT 3

US-09-621-976-3915  
Sequence 3915, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 3915  
LENGTH: 95  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -18...-1

NAME/KEY: UNSURE  
LOCATION: 59  
OTHER INFORMATION: Xaa = Ala,Thr  
NAME/KEY: UNSURE  
LOCATION: 56  
OTHER INFORMATION: Xaa = Arg,Ser  
NAME/KEY: UNSURE  
LOCATION: 54  
OTHER INFORMATION: Xaa = His,Pro  
US-09-621-976-3915

Query Match 30.4%; Score 494; DB 4; Length 95;  
Best Local Similarity 96.8%; Pred. No. 1.7e-46;  
Matches 92; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAVYILLCWAGLPICLATCDLHPTGSRPTVPGLHPSGY 60  
Db 1 MKAPGRVLIIICSVFSAVYILLCWAGLPICLATCDLHPTGSRPTVPGLHPSGY 60  
QY 61 SVDPGKEIVREPCRSQAVVSSSGOMLGSLGAEID 95  
Db 61 SVDPGKEIVREPCRSQAVVSSSGOMLGSLGAEID 95

## RESULT 4

US-08-446-875-16  
Sequence 16, Application US/08446875  
Patent No. 5858751

GENERAL INFORMATION:  
APPLICANT: Paulson, James C.  
APPLICANT: Wen, Xiaohong  
APPLICANT: Livingston, Brian Duane  
APPLICANT: Gillespie, William  
APPLICANT: Keim, Jorge  
APPLICANT: Burlingame, Alma L.  
TITLE OF INVENTION: Medzimbazsky, Katalin  
TITLE OF INVENTION: Compositions and Methods for the  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Poms, Smith, Lande & Rose  
STREET: 2029 Century Park East, Suite 3800  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90067

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,875  
FILING DATE: July 12, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/102,385  
FILING DATE: August 4, 1993

ATTORNEY/AGENT INFORMATION:  
NAME: Oldenkamp, David J.  
REGISTRATION NUMBER: 29,421  
REFERENCE/DOCKET NUMBER: 111-197  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (310) 788-5000  
TELEFAX: (310) 277-1297  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-446-875-16

Query Match 12.2%; Score 198.5; DB 2; Length 340;  
Best Local Similarity 27.8%; Pred. No. 3.6e-13;  
Matches 55; Conservative 32; Mismatches 60; Indels 51; Gaps 6;

QY 73 CRCAVSSSGOMLGSGLAIDSAECVFRMNOAPTVGEADVQGRSTLRVSHSTVPL 132  
DB 139 CRCAVSSSGOMLGSGLAIDSAECVFRMNOAPTVGEADVQGRSTLRVSHSTVPL 130  
QY 133 LKNYSYFQKARDTLVYVWGQGRHMD-----RVLGRTYRTLL-----QLTRM 175  
DB 191 ---YPSFRELGDVNSMILVFPKIDLEWVSAITTGTSHTYIYPVPAKIRVQDKILY 247  
QY 176 YPGLQVYTFERRMAYCDQIFQDETGKRGSGSFLSTGFWMTIMALECEIIVYGMS 235  
DB 248 HPAFIKTVFQNMVQGH-----GRYPTGILSVIFSMHVCDEVLDYGFGA 291

QY 236 DSYCREKSHPSVPYHYFE 253  
DB 292 DS--KGNMH-----HYWE 302

RESULT 5  
US-08-503-133A-2  
; Sequence 2, Application US/08503133A  
; Patent No. 5747326  
; GENERAL INFORMATION:  
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;  
; TITLE OF INVENTION: Isolated Polyalyl Transferases,  
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of  
; TITLE OF INVENTION: Production and Use  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felle & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/503,133A  
; FILING DATE: 17-JULY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/04289  
; FILING DATE: 22-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5747326man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: BOER 1050  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-503-133A-2

Query Match 12.1%; Score 196.5; DB 1; Length 359;  
Best Local Similarity 27.5%; Pred. No. 6.4e-13;  
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPQKPLVREPCSCAVSSSGOMLGSGLAIDSAECVFRMNOAPTVGEADVQGR 118  
DB 125 HSLLEPVSPMKRRRFTKCAVVGNSGILLDSGCKEIDSHNFVIRCLAPVVEFADVGT 164

QY 119 STRVSVSHSVPL---LKNYS-----HYFQARDTL-----YVWGGRHMDRVYGR 164  
DB 185 SDFITMNPVSVOQAFGCFRNSIDRAKVFHRLSMINDSVLWIPAFMVGSGKHVEW-NAL 243  
QY 165 TYRTLLQLTMYGLOV-----YTFERRMAYCDQIFQDETGKRGSGSFLSTGFWMT 218  
DB 244 ILNKLKVRTAYPSLRIHAVRGIMLNKVP-----KRP-----STGLLMY 285

QY 219 IIALLECEIIVYGMSDSYCREKSHPSVPYHYFEKGRLECOMVLAHQAPPSAHRFIT 278  
DB 286 TLATRFQDEIHLGFW--PPKDLNGKAVKXHYD---DLKRYTSNAPSHPMPLEFKT 339

QY 279 EKAVFSRMAK 289  
DB 340 LVNLIHNRGALK 350

RESULT 6  
US-08-576-775A-2  
; Sequence 2, Application US/08576775A  
; Patent No. 5849904  
; GENERAL INFORMATION:  
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;  
; TITLE OF INVENTION: Isolated Polyalyl Transferases,  
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of  
; TITLE OF INVENTION: Production and Use  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felle & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/576,775A  
; FILING DATE: 21-December-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/503,133  
; FILING DATE: 17-July-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/04289  
; FILING DATE: 22-December-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5849904man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: BOER 1050.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-576-775A-2

Query Match 12.1%; Score 196.5; DB 2; Length 359;  
Best Local Similarity 27.5%; Pred. No. 6.4e-13;  
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPQKPLVREPCSCAVSSSGOMLGSGLAIDSAECVFRMNOAPTVGEADVQGR 118  
DB 125 HSLLEPVSPMKRRRFTKCAVVGNSGILLDSGCKEIDSHNFVIRCLAPVVEFADVGT 164



```

QY      113  SGLTAVSHSHSVPLT---LENYG-----HYEQKADLT-----YMWGQGGSHDRVYIGR 164
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      185  SDFITNPSVYQARFEGCFRNESDRAKFVHRUSLNDSTVLTLPAMVKGGEKHEWV-NAL 243
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      165  TYRTLLQLTFRMYEGLQV-----YTETERMAVYCDQIFODETGKRRQSGFLSTGFTM 218
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      244  ILNKKIKVTAAYSLRLIHAIVGMYLTNNKVP-----KRP-----STGLLMY 285
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      219  IIALLECEELVYVGWVSDSYCREKSHPSVYPERKGLDECOMYTLAEQA.PRAHHFIT 278
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      286  TLTATRCDEIHLTGFW--PFPKDLNGKAVKXHYVD---DKYRFYSNASPHRMPLFEKT 339
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      279  EKAVFSRMAKX 289
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      340  LNVLTNNKGALK 350
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

1      RESULT 7
2      US-08-972-498-2
3      Sequence 2, Application US/08972498
4      Patent No. 595078
5      GENERAL INFORMATION:
6      APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
7      APPLICANT: Nakayama, Jun; Eckhardt, Mathias
8      TITLE OF INVENTION: Isolated Polysialyl Transferases,
9      TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
10     TITLE OF INVENTION: Production and Use
11     NUMBER OF SEQUENCES: 8
12     CORRESPONDENCE ADDRESS:
13     ADDRESSEE: Felte & Lynch
14     STREET: 805 Third Avenue
15     CITY: New York City
16     STATE: New York
17     COUNTRY: USA
18     ZIP: 10022
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
21     COMPUTER: IBM PS/2
22     OPERATING SYSTEM: PC-DOS
23     SOFTWARE: Wordperfect
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/972,498
26     FILING DATE:
27     CLASSIFICATION: 536
28     PRIOR APPLICATION DATA:
29     APPLICATION NUMBER: 08/576,775
30     FILING DATE: 21-December-1995
31     APPLICATION NUMBER: 08/503,133
32     FILING DATE: 17-July-1995
33     PRIOR APPLICATION DATA:
34     APPLICATION NUMBER: PCT/EP94/04289
35     FILING DATE: 22-December-1994
36     ATTORNEY/AGENT INFORMATION:
37     NAME: Hanson, No. 5959078man D.
38     REGISTRATION NUMBER: 30,946
39     REFERENCE/DOCKET NUMBER: BOER 1050.1
40     TELECOMMUNICATION INFORMATION:
41     TELEPHONE: (212) 688-9200
42     TELEFAX: (212) 838-3884
43     INFORMATION FOR SEQ ID NO: 2:
44     SEQUENCE CHARACTERISTICS:
45     LENGTH: 359 amino acids
46     TYPE: amino acid
47     TOPOLOGY: linear
48     MOLECULE TYPE: protein
49     US-08-972-498-2

```

Query Match	12.1%;	Score 196.5;	DB 2;	Length 359;
Best Local Similarity	27.5%;	Pred. No. 6.4e-13;		
Matches	69;	Conservative	39;	Mismatches 98; Indels 45; Gaps 9
QY	59	YSSVDPGKLVAPPCSCAVVSSSGQMGSGIGAEITDSAEVFRNMQAPTVGEADVQR	118	

Db 125 HSLTPEVSMKRRPKRTCAVGNISGILLDSCGKEIDSHNFVLRICNLAPVFEFAADVGTK 164  
QY 119 SLLRVVSHTSVPL--LRNYS-----HYQKARDL-----YMWGQGRHMDRYLGR 164  
Db 185 SDFITNPSVVRPAFGFGRNESDRAKFVHRLSTMLSDVLMIPAFMVYGGKHEWV-NAL 243  
QY 165 YKRTLLQLTRMYPGLQV-----YTFERMAVYCDQIFQDETGNRRQSSFLSTGFTM 218  
Db 244 ILNKRKLKVTAAPSRLTHAVAGVWLTNKVLI-----KRP-----STGLMT 285  
QY 219 IIALLECEIIVYGVMSDYCRKSHPSVPVHYFEKRLDECOMYLAEQA PRSAREFT 278  
Db 286 TLATRRCDIEHLXGFW--PFRDLNGKAVKHYVD---DLKRYVFSNAPSRHMRPLEFKT 339  
QY 279 EKAVFSRMAKK 289  
Db 340 LNVVHNRGALX 350

```

1      RESULT 8
2      US-08-899-545-2
3      Sequence 2, Application US/08893545
4      Patent No. 6020201
5      GENERAL INFORMATION:
6      APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
7      APPLICANT: Nakayama, Jun; Eckhardt, Matthias
8      TITLE OF INVENTION: Isolated Polysialyl Transferase,
9      TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
10     NUMBER OF SEQUENCES: 8
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: Felte & Lynch
13     STREET: 805 Third Avenue
14     CITY: New York City
15     STATE: New York
16     COUNTRY: USA
17     ZIP: 10022
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
20     COMPUTER: IBM PS/2
21     OPERATING SYSTEM: PC-DOS
22     SOFTWARE: Wordperfect
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/899,545
25     FILING DATE:
26     CLASSIFICATION:
27     PRIOR APPLICATION DATA:
28     APPLICATION NUMBER: 08/503,133
29     FILING DATE:
30     ATTORNEY/AGENT INFORMATION:
31     NAME: Hanson, No. 6020201man D.
32     REGISTRATION NUMBER: 30,946
33     REFERENCE/DOCKET NUMBER: BOER 1050
34     TELECOMMUNICATION INFORMATION:
35     TELEPHONE: (212) 688-9200
36     TELEFAX: (212) 838-3884
37     INFORMATION FOR SEQ ID NO: 2:
38     SEQUENCE CHARACTERISTICS:
39     LENGTH: 359 amino acids
40     TYPE: amino acid
41     TOPOLOGY: linear
42     MOLECULE TYPE: protein
43     US-08-899-545-2

```

```

Query Match 12.1%; Score 196.5; DB 3; length 359;
Best Local Similarity 27.5%; Pred. No. 6.4e-13;
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9

QY 59 YSSVDGKPLVREPRSCANVSSGCGMAGIGATIDDAECFFRNNOAPTVFEEADVGR 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 HSLLEVPKMKRRRTKCAVVGNSGILIDSGGKEIDSNFVIRCLNAPVFEFADVGK 184

```

```

QY      119  STLRVSTSVPLT-----RNRYS-----HYFOKARFTL-----YMWGCOHRHDIYVGG  164
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      185  SDFITMNSVYQORAGGFRNESDRAKFYHRLSMINDSVLMIRAFMWKGGKHEVNW--NAL  243
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      165  TYRFTLLQLTRMYPGLOV-----YFTTERMMAYCDOI FODETGKNNRQGGSPFSTGMPFTM  218
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      244  ILKNKLLKRTAYVPSLRILHAIVRGWMLTNKVP I-----XRP-----STGLLMY  285
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      219  ILALELCEIIVYVGNVSDSYCREKSHSPYHYFEKGRLDCEOMTYLAHEOAPRSARFIT  278
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      286  TLAIRRFCEIHLVGFW--PFPDRLNGAKVKNHYD--DLKRYSPSNASPHRMPLDEFT  339
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      279  EKAVERWAKK  289
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      340  LNVLTNRCALK  350
           :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

```

1      RESULT 9
2      US-08-446-875-2
3      / Sequence 2, Application US/08446875
4      / Patent No. 5858751
5      /
6      / GENERAL INFORMATION:
7      / APPLICANT: Paulson, James C.
8      / APPLICANT: Wen, Xiaohong
9      / APPLICANT: Livingston, Brian Duane
10     / APPLICANT: Gillespie, William
11     / APPLICANT: Kelm, Serge
12     / APPLICANT: Burlingame, Alma L.
13     / APPLICANT: Medzihradzky, Katalin
14     / TITLE OF INVENTION: Compositions and Methods for the
15     / TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
16     / NUMBER OF SEQUENCES: 16
17     /
18     / CORRESPONDENCE ADDRESSES:
19     / ADDRESSEE: Potts, Smith, Lande & Rose
20     / STREET: 2029 Century Park East, Suite 3800
21     / CITY: Los Angeles
22     / STATE: CA
23     / COUNTRY: USA
24     /
25     / ZIP: 90067
26     /
27     / COMPUTER READABLE FORM:
28     / MEDIUM TYPE: Floppy disk
29     / COMPUTER: IBM PC compatible
30     / OPERATING SYSTEM: PC-DOS/MS-DOS
31     / SOFTWARE: PatentIn Release #1.0, Version #1.25
32     /
33     / CURRENT APPLICATION DATA:
34     / APPLICATION NUMBER: US/08/446,875
35     / FILING DATE: July 12, 1995
36     / CLASSIFICATION: 435
37     /
38     / PRIOR APPLICATION DATA:
39     / APPLICATION NUMBER: 08/102,385
40     / FILING DATE: August 4, 1993
41     /
42     / ATTORNEY/AGENT INFORMATION:
43     / NAME: Oldenkamp, David J.
44     / REGISTRATION NUMBER: 29,421
45     / REFERENCE/DOCKET NUMBER: 111-197
46     /
47     / TELECOMMUNICATION INFORMATION:
48     / TELEPHONE: (310) 788-5000
49     / TELEFAX: (310) 277-1297
50     /
51     / INFORMATION FOR SEQ ID NO: 2:
52     / SEQUENCE CHARACTERISTICS:
53     / LENGTH: 343 amino acids
54     / TYPE: amino acid
55     / TOPOLOGY: linear
56     /
57     / MOLECULE TYPE: protein
58     /
59     /
60     /
61     /
62     /
63     /
64     /
65     /
66     /
67     /
68     /
69     /
70     /
71     /
72     /
73     /
74     /
75     /
76     /
77     /
78     /
79     /
80     /
81     /
82     /
83     /
84     /
85     /
86     /
87     /
88     /
89     /
90     /
91     /
92     /
93     /
94     /
95     /
96     /
97     /
98     /
99     /
100    /
101    /
102    /
103    /
104    /
105    /
106    /
107    /
108    /
109    /
110    /
111    /
112    /
113    /
114    /
115    /
116    /
117    /
118    /
119    /
120    /
121    /
122    /
123    /
124    /
125    /
126    /
127    /
128    /
129    /
130    /
131    /
132    /
133    /
134    /
135    /
136    /
137    /
138    /
139    /
140    /
141    /
142    /
143    /
144    /
145    /
146    /
147    /
148    /
149    /
150    /
151    /
152    /
153    /
154    /
155    /
156    /
157    /
158    /
159    /
160    /
161    /
162    /
163    /
164    /
165    /
166    /
167    /
168    /
169    /
170    /
171    /
172    /
173    /
174    /
175    /
176    /
177    /
178    /
179    /
180    /
181    /
182    /
183    /
184    /
185    /
186    /
187    /
188    /
189    /
190    /
191    /
192    /
193    /
194    /
195    /
196    /
197    /
198    /
199    /
200    /
201    /
202    /
203    /
204    /
205    /
206    /
207    /
208    /
209    /
210    /
211    /
212    /
213    /
214    /
215    /
216    /
217    /
218    /
219    /
220    /
221    /
222    /
223    /
224    /
225    /
226    /
227    /
228    /
229    /
230    /
231    /
232    /
233    /
234    /
235    /
236    /
237    /
238    /
239    /
240    /
241    /
242    /
243    /
244    /
245    /
246    /
247    /
248    /
249    /
250    /
251    /
252    /
253    /
254    /
255    /
256    /
257    /
258    /
259    /
260    /
261    /
262    /
263    /
264    /
265    /
266    /
267    /
268    /
269    /
270    /
271    /
272    /
273    /
274    /
275    /
276    /
277    /
278    /
279    /
280    /
281    /
282    /
283    /
284    /
285    /
286    /
287    /
288    /
289    /
290    /
291    /
292    /
293    /
294    /
295    /
296    /
297    /
298    /
299    /
300    /
301    /
302    /
303    /
304    /
305    /
306    /
307    /
308    /
309    /
310    /
311    /
312    /
313    /
314    /
315    /
316    /
317    /
318    /
319    /
320    /
321    /
322    /
323    /
324    /
325    /
326    /
327    /
328    /
329    /
330    /
331    /
332    /
333    /
334    /
335    /
336    /
337    /
338    /
339    /
340    /
341    /
342    /
343    /
344    /
345    /
346    /
347    /
348    /
349    /
350    /
351    /
352    /
353    /
354    /
355    /
356    /
357    /
358    /
359    /
360    /
361    /
362    /
363    /
364    /
365    /
366    /
367    /
368    /
369    /
370    /
371    /
372    /
373    /
374    /
375    /
376    /
377    /
378    /
379    /
380    /
381    /
382    /
383    /
384    /
385    /
386    /
387    /
388    /
389    /
390    /
391    /
392    /
393    /
394    /
395    /
396    /
397    /
398    /
399    /
400    /
401    /
402    /
403    /
404    /
405    /
406    /
407    /
408    /
409    /
410    /
411    /
412    /
413    /
414    /
415    /
416    /
417    /
418    /
419    /
420    /
421    /
422    /
423    /
424    /
425    /
426    /
427    /
428    /
429    /
430    /
431    /
432    /
433    /
434    /
435    /
436    /
437    /
438    /
439    /
440    /
441    /
442    /
443    /
444    /
445    /
446    /
447    /
448    /
449    /
450    /
451    /
452    /
453    /
454    /
455    /
456    /
457    /
458    /
459    /
460    /
461    /
462    /
463    /
464    /
465    /
466    /
467    /
468    /
469    /
470    /
471    /
472    /
473    /
474    /
475    /
476    /
477    /
478    /
479    /
480    /
481    /
482    /
483    /
484    /
485    /
486    /
487    /
488    /
489    /
490    /
491    /
492    /
493    /
494    /
495    /
496    /
497    /
498    /
499    /
500    /
501    /
502    /
503    /
504    /
505    /
506    /
507    /
508    /
509    /
510    /
511    /
512    /
513    /
514    /
515    /
516    /
517    /
518    /
519    /
520    /
521    /
522    /
523    /
524    /
525    /
526    /
527    /
528    /
529    /
530    /
531    /
532    /
533    /
534    /
535    /
536    /
537    /
538    /
539    /
540    /
541    /
542    /
543    /
544    /
545    /
546    /
547    /
548    /
549    /
550    /
551    /
552    /
553    /
554    /
555    /
556    /
557    /
558    /
559    /
560    /
561    /
562    /
563    /
564    /
565    /
566    /
567    /
568    /
569    /
570    /
571    /
572    /
573    /
574    /
575    /
576    /
577    /
578    /
579    /
580    /
581    /
582    /
583    /
584    /
585    /
586    /
587    /
588    /
589    /
590    /
591    /
592    /
593    /
594    /
595    /
596    /
597    /
598    /
599    /
600    /
601    /
602    /
603    /
604    /
605    /
606    /
607    /
608    /
609    /
610    /
611    /
612    /
613    /
614    /
615    /
616    /
617    /
618    /
619    /
620    /
621    /
622    /
623    /
624    /
625    /
626    /
6
```

[illegible]

```

QY      124 VSHNSVPLLENNNSYFQKARDLTLMWGGGRHMRVLGGRTYRTLLDQLTRMYPGLQVYT 183
Db      193 YYPBSFRELAQVS-----MLIVPKTTDLELVISATTTGIIISHYVPPA--KIKY 242
QY      184 FTERMAVACDDIQFODETGNRRQSGSPSTGWFMTMLALELCEBIWVGKSPDSYCREKS 243
Db      243 KKEKLLIHPAFIKIVFDPRMLQGHGRYRSTGLISVIFSLHICDEVDLYGFGADS--KGNW 300
QY      244 HPSVPHYHYS 253
Db      301 H-----HYWE 305

```

```

10 RESULT 10
11 US-08-102-385G-2
12 ; Sequence 2, Application US/08102385G
13 ; Patent No. 5962294
14 ; GENERAL INFORMATION:
15 ; APPLICANT: Paulson, James C.
16 ; APPLICANT: Wen, Xiaohong
17 ; APPLICANT: Livingston, Brian Duane
18 ; APPLICANT: Gillespie, William
19 ; APPLICANT: Kelm, Sotge
20 ; APPLICANT: Burlingame, Alma L.
21 ; APPLICANT: Medzhiradzsky, Katalin
22 ; TITLE OF INVENTION: Compositions and Methods for the
23 ; TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
24 ; NUMBER OF SEQUENCES: 36
25 ; CORRESPONDENCE ADDRESS:
26 ; ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
27 ; STREET: 2029 Century Park East, 38th Floor
28 ; CITY: Los Angeles
29 ; STATE: CA
30 ; COUNTRY: USA
31 ; ZIP: 90067
32 ; COMPUTER READABLE FORM:
33 ; MEDIUM TYPE: Floppy disk
34 ; COMPUTER: IBM PC compatible
35 ; OPERATING SYSTEM: PC-DOS/MS-DOS
36 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
37 ; CURRENT APPLICATION DATA:
38 ; APPLICATION NUMBER: US/08/102.385G
39 ; FILING DATE: 04-AUG-1993
40 ; CLASSIFICATION: 435
41 ; PRIOR APPLICATION DATA:
42 ; APPLICATION NUMBER: US 07/925369
43 ; FILING DATE: 04-AUG-1992
44 ; ATTORNEY/AGENT INFORMATION:
45 ; NAME: Oldenkamp, David J.
46 ; REGISTRATION NUMBER: 29,421
47 ; REFERENCE/DOCKET NUMBER: 97-062
48 ; TELECOMMUNICATION INFORMATION:
49 ; TELEPHONE: (310)788-5000
50 ; TELEFAX: (310) 277-1297
51 ; INFORMATION FOR SEQ ID NO: 2:
52 ; SEQUENCE CHARACTERISTICS:
53 ; LENGTH: 343 amino acids
54 ; TYPE: amino acid
55 ; TOPOLOGY: linear
56 ; MOLECULE TYPE: protein
57 ; US-08-102-385G-2

```

```

Query Match      11.7%: Score 191; DB 2; Length 343;
Best Local Similarity 30.0%: Pred. No. 2.4e-12;
Matches 57; Conservative 36; Mismatches 77; Indels 20; Gaps 5

QY 67 PLVRE---PCRSQAVSSSGQMLTSGGAEIDSABCFVRNMQAFTVGSEADVGQRSTURV 123
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 133 PLERELVLSRCRCVAVGNISGNLTKESYGPQIDSHDFVLRRNNKAPTEGSEADVGSKTHHF 192
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 124 VSHTSVPLLENNYSHYFOKADDTLMTMGQGRHMDRVLGSGTYRTLLDTMTMPQLQVYT 183
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 193 VYSESPRELAQEV-----MILVFKTDLLEWISATTTGTISHTVVPVA-KIKV 242

QY 184 FTRMMAYCDQIFQDETKNRROSGSLSTGWTMLALCEIIVYGWVSSYCKRS 243

Db 243 KKEKILYHAFIKYVDFRWLOCHGRYPSTGILSVIFSLHCDEVDLYFGADS--KGNW 300

QY 244 HPSVPYHYFE 253

Db 301 H-----HYWE 305

## RESULT 11

US-08-626-994A-3

; Sequence 3, Application US/08626994A

; Patent No. 5798244

; GENERAL INFORMATION:

; APPLICANT: Shuichi TSUJI et al.

; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ' 2,8-

; TITLE OF INVENTION: SIATYLTRANSFERASE

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenblum & Bernstein, P.L.C.

; STREET: 1941 Roland Clarke Place

; City: Reston

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 20191

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 diskette

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Word Perfect 5.1+ (ASCII)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/626,994A

; FILING DATE: April 3, 1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: JP 77469/1995

; FILING DATE: April 3, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Arnold Turk

; REGISTRATION NUMBER: 33,094

; REFERENCE/DOCKET NUMBER: P14595

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 716-1191

; TELEFAX: (703) 716-1180

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 339 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

US-08-626-994A-3

Query Match 11.7%; Score 190.5; DB 1; Length 339;

Best Local Similarity 30.0%; Pred. No. 2,7e-12;

Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGRPLVREPCRCGAVVSSGQMLGSLGAEIDSACEVFRMNOAPTVGFADVGORS 119

Db 105 SLIPDVSPINMKRYNCAVVGNSGILTGSCQGEIDKSDVSRCNFAPTEAFHKDVGRT 164

QY 120 TLKRVSHTSVPLLRNYSHFQKARDTL-----YMWGQGHMDRVLGRTYRT 168

Db 165 NLT-----TNPSTILEKYYNMLTIQDRNNFFLSLKLGDALW-----IPAFFHT 211

QY 169 LLQLTRMYPGLQYTYFTER-----MMAYCDQIFQ--DETKNRROSGSLSTGWTML 220

Db 212 SATVTRTL-----VDFVEHNGQLKVQLAMPGNIMQHNRYKMKHLSPKRLSTGILMYTL 267

QY 221 ALCEIIVVYGWVSS--DSYCKEKSHPSVPHYHYFEK 254

Db 268 ASAIICEIHLXGFWPFGFDPNTR-----DLPYHYDK 300

## RESULT 12

US-08-957-742-3

; Sequence 3, Application US/08957742

; Patent No. 6017743

; GENERAL INFORMATION:

; APPLICANT: Shuichi TSUJI et al.

; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ' 2,8-

; TITLE OF INVENTION: SIATYLTRANSFERASE

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenblum & Bernstein, P.L.C.

; STREET: 1941 Roland Clarke Place

; City: Reston

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 20191

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 diskette

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Word Perfect 5.1+ (ASCII)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/957,742

; FILING DATE:

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/626,994

; FILING DATE: April 3, 1996

; APPLICATION NUMBER: JP 77469/1995

; FILING DATE: April 3, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Arnold Turk

; REGISTRATION NUMBER: 33,094

; REFERENCE/DOCKET NUMBER: P14595

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 716-1191

; TELEFAX: (703) 716-1180

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 339 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

US-08-957-742-3

Query Match 11.7%; Score 190.5; DB 3; Length 339;

Best Local Similarity 30.0%; Pred. No. 2,7e-12;

Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGRPLVREPCRCGAVVSSGQMLGSLGAEIDSACEVFRMNOAPTVGFADVGORS 119

Db 105 SLIPDVSPINMKRYNCAVVGNSGILTGSCQGEIDKSDVSRCNFAPTEAFHKDVGRT 164

QY 120 TLKRVSHTSVPLLRNYSHFQKARDTL-----YMWGQGHMDRVLGRTYRT 168

Db 165 NLT-----TNPSTILEKYYNMLTIQDRNNFFLSLKLGDALW-----IPAFFHT 211

QY 169 LLQLTRMYPGLQYTYFTER-----MMAYCDQIFQ--DETKNRROSGSLSTGWTML 220

Db 212 SATVTRTL-----VDFVEHNGQLKVQLAMPGNIMQHNRYKMKHLSPKRLSTGILMYTL 267

QY 221 ALCEIIVVYGWVSS--DSYCKEKSHPSVPHYHYFEK 254

Db 268 ASAIICEIHLXGFWPFGFDPNTR-----DLPYHYDK 300

## RESULT 13

US-08-626-994A-1

; Sequence 1, Application US/08626994A

; Patent No. 5798244

; GENERAL INFORMATION:

```

? COUNTRY: U.S.A.
? ZIP: 20191
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 1.44 diskette
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: Word Perfect 5.1+ (ASCII)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/957, 742
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/626, 994
? FILING DATE: April 3, 1996
? APPLICATION NUMBER: JP 77469/1995
? FILING DATE: April 3, 1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Arnold Turk
? REGISTRATION NUMBER: 33,094
? REFERENCE/DOCKET NUMBER: P14595
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 716-1191
? TELEFAX: (703) 716-1180
? INFORMATION FOR SEQ ID NO: 1 :
? SEQUENCE CHARACTERISTICS:
? LENGTH: 364 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: linear
US-08-957-742-1
Query Match          11.7%; Score 190.5; DB 3; Length 364;
Best Local Similarity 30.0%; Pred.No. 3e-12;
Matches    65; Conservative   28; Mismatches   81; Indels   43; Gaps    8;

QY      60 SSVDPGKELVREPCSGAVSSGGOMLGSGLGAETDSEACVFRRNQAPTVGFEADVQRS 119
        |:|:::~::~|||:|||:::|||||:::|||||:::|||||:::|||||:::
DB       130 SLPEVPSPINKRKNRYCAVANGSCLTGSOGGOEIDKSDPVSRCNFAPTEAFHKVGRKT 189
               |||:::~::~|||:|||:::|||||:::|||||:::|||||:::
QY      120 TLRVSTSYVLILRNYSHYFOKRDTL-----IYWVGCGRHMDRVLGRTYRT 168
        | 190 NLT---TFNPSTLEKYNNLLTIQDNKFFLSLKDGNALW-----IPAFFHT 236
               |||:::~::~|||:|||:::|||||:::|||||:::|||||:::
DB       237 SATVTRTL----VDFFVEHRGOLKVQLAMPGINIWGHNRVKMKHLSPKRLSTGIWLTYTL 292
               |||:::~::~|||:|||:::|||||:::|||||:::|||||:::
QY      169 LLOITRMYPGLQVVYTFTER-----MMAYCQQIFQ--DETCKQRSGSFSLTGWFMTML 220
        | 221 ALBELCEIVYYGMWS---DSYCERKSHPSPFYHFEEK 254
        |         ::|||:::~::~|||:|||:::|||||:::|||||:::
DB       293 ASAICEIHLYGFWPFGFDPTBRE---DLPHYDYDK 325
               |||:::~::~|||:|||:::|||||:::|||||:::|||||:::

RESULT 15
US-08-666-367B-6
Sequence 6, Application US/08666367B
Patent No. 5854042
GENERAL INFORMATION:
APPLICANT: Shuichi TSUJI et al.
TITLE OF INVENTION: NOVEL SUGAR-CHAIN SYNTHETASE AND PROCESS FOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSES: Wenderoch, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: wordperfect 5.1
```

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/666,367B  
FILING DATE: August 19, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 355 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-666-367B-6

Query Match 10.9%; Score 178; DB 2; Length 355;  
Best Local Similarity 25.2%; Pred. No. 6.8e-11;  
Matches 63; Conservative 37; Mismatches 80; Indels 70; Gaps 9;  
QY 73 CRGAVVSSGQMLGSLGAEIDSAECVPMNOAPTVGFADVGQSTLRVSHTSVPLL 132  
DB 126 CISCAYVNGGILNNSGMEIDSHDYFRVSGAVIKGIEKDVGTSTFGFTAYSLVSS 185  
QY 133 LRNSHY-FOKAPDLYVMWGQGRHMDRVLGRTYRTLLQITRNYPGIQQV----- 181  
DB 186 LQNGHKGFKI-----PQKHI-----RYHFLKAVRDYEWIKALLDKDIRKG 230  
QY 182 -----YTFTRMMAYCDQIFQDETGNRRQSSFL-----STGW 215  
DB 231 FLNYYGRPRERFDEFTMNKYLVAAHDPFL--RYLKNRFLKSKNLOKPYWRLYRPTTGA 287  
QY 216 FTMLALELCEIIVYGVSDSYCREKSHPSVPYHYFEKGRLDCEOMYLAHQAPRSAMR 275  
DB 288 LLLITALLHLCDRVSAHYGITE-----GHQKYSDHYVDK-EWKRLVVFVYVNH----- 332  
QY 276 FITEKAVPSR 285  
DB 333 FNLEKQVWKR 342

Search completed: December 21, 2004, 22:44:59  
Job time : 27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 22:13:05 ; Search time 77 Seconds

(without alignments)  
1403.457 Million cell updates/sec

Title: US-10-019-114A-7

Perfect score: 1626

Sequence: 1 MKRPGRLVLIILCSVFSVSAV.....FSRWAKRPVFAHPSWRTS 302

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	100.0	302	11	US-09-978-360A-407
2	1626	100.0	302	14	US-10-315-664-75
3	1622	99.8	302	10	US-09-895-298-66
4	1622	99.8	302	10	US-09-895-298-120
5	1167	71.8	218	15	US-10-262-511-184
6	726.5	44.7	305	17	US-10-844-874-20
7	716.5	44.1	305	11	US-09-833-245-1914
8	716.5	44.1	305	11	US-09-833-245-1915
9	714.5	43.9	305	15	US-10-188-186-22
10	713.5	43.9	305	14	US-10-104-047-2350
11	694	42.7	318	15	US-10-188-186-24
12	677.5	41.7	305	15	US-10-188-186-20
13	639.5	39.3	336	14	US-10-097-340-189

14	639.5	39.3	336	14	US-10-094-749-2441	Sequence 2441, Ap
15	636.5	39.1	336	14	US-10-288-252-11	Sequence 11, Appl
16	636.5	39.1	336	10	US-10-398-038-11	Sequence 11, Appl
17	537	33.0	299	10	US-09-946-374-56	Sequence 56, Appl
18	537	33.0	299	13	US-10-052-586-206	Sequence 206, App
19	537	33.0	299	14	US-10-174-590-206	Sequence 206, App
20	537	33.0	299	14	US-10-176-758-206	Sequence 206, App
21	537	33.0	299	14	US-10-175-737-206	Sequence 206, App
22	537	33.0	299	14	US-10-174-581-206	Sequence 206, App
23	537	33.0	299	14	US-10-176-483-206	Sequence 206, App
24	537	33.0	299	14	US-10-176-749-206	Sequence 206, App
25	537	33.0	299	14	US-10-176-914-206	Sequence 206, App
26	537	33.0	299	14	US-10-176-915-206	Sequence 206, App
27	537	33.0	299	14	US-10-173-706-206	Sequence 206, App
28	537	33.0	299	14	US-10-175-738-206	Sequence 206, App
29	537	33.0	299	14	US-10-175-752-206	Sequence 206, App
30	537	33.0	299	14	US-10-176-482-206	Sequence 206, App
31	537	33.0	299	14	US-10-176-757-206	Sequence 206, App
32	537	33.0	299	14	US-10-176-913-206	Sequence 206, App
33	537	33.0	299	14	US-10-180-552-206	Sequence 206, App
34	537	33.0	299	14	US-10-180-553-206	Sequence 206, App
35	537	33.0	299	14	US-10-173-700-206	Sequence 206, App
36	537	33.0	299	14	US-10-174-572-206	Sequence 206, App
37	537	33.0	299	14	US-10-174-573-206	Sequence 206, App
38	537	33.0	299	14	US-10-174-582-206	Sequence 206, App
39	537	33.0	299	14	US-10-174-588-206	Sequence 206, App
40	537	33.0	299	14	US-10-175-739-206	Sequence 206, App
41	537	33.0	299	14	US-10-175-740-206	Sequence 206, App
42	537	33.0	299	14	US-10-175-743-206	Sequence 206, App
43	537	33.0	299	14	US-10-176-488-206	Sequence 206, App
44	537	33.0	299	14	US-10-176-492-206	Sequence 206, App
45	537	33.0	299	14	US-10-176-747-206	Sequence 206, App

#### ALIGNMENTS

RESULT 1  
US-09-978-360A-407  
; Sequence 407, Application US/09978360A  
; Publication No. US20040110939A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
; APPLICANT: Ducleret, Aymric  
; APPLICANT: Bouguetieret, Lydie  
; APPLICANT: Joberet, Severin  
; APPLICANT: Clusel, Catherine  
; TITLE OR INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: 56.USA.CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: US 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: US 60/099,273  
; PRIOR FILING DATE: -09-04  
; PRIOR APPLICATION NUMBER: US 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: US 09/215,435  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: PCT/IB98/02122  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: US 09/247,155  
; PRIOR FILING DATE: 1999-02-09  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 810

SOFTWARE: Patent.pm  
SEQ ID NO 407  
LENGTH: 302  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -18...-1  
US-09-978-360A-407

Query Match 100.0%; Score 1626; DB 11; Length 302;  
Best Local Similarity 100.0%; Pred. No. 2.4e-160;  
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVPGPLHFSGYG 60  
DB 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVPGPLHFSGYG 60  
QY 61 SVDPGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAFTVGFADVGQRST 120  
DB 61 SVDPGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAFTVGFADVGQRST 120  
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLQLTRMYPGLQ 180  
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLQLTRMYPGLQ 180  
QY 181 VYTFTERMAYCDQIFODETGKRRQSGFLSTGFTMLALELCEIIVYGVMSDSYCR 240  
DB 181 VYTFTERMAYCDQIFODETGKRRQSGFLSTGFTMLALELCEIIVYGVMSDSYCR 240  
QY 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300  
DB 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300  
QY 301 TE 302  
DB 301 TE 302

## RESULT 2

US-10-315-664-75  
Sequence 75, Application US/10315664  
Publication No. US20030203377A1  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Bouguetieret, L.  
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal  
FILE REFERENCE: GENSET.050CP3  
CURRENT APPLICATION NUMBER: US/10/315,664  
PRIOR FILING DATE: 2002-12-09  
PRIOR APPLICATION NUMBER: US/09/599,360  
PRIOR FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: 60/113,686  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/141,032  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 09/469,099  
PRIOR FILING DATE: 1999-12-21  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: Patent.pm  
SEQ ID NO 75  
LENGTH: 302  
TYPE: PRT  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -18...-1  
US-10-315-664-75

Query Match 100.0%; Score 1626; DB 14; Length 302;  
Best Local Similarity 100.0%; Pred. No. 2.4e-160;

Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVPGPLHFSGYG 60  
DB 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVPGPLHFSGYG 60  
QY 61 SVDPGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAFTVGFADVGQRST 120  
DB 61 SVDPGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAFTVGFADVGQRST 120  
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLQLTRMYPGLQ 180  
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLQLTRMYPGLQ 180  
QY 181 VYTFTERMAYCDQIFODETGKRRQSGFLSTGFTMLALELCEIIVYGVMSDSYCR 240  
DB 181 VYTFTERMAYCDQIFODETGKRRQSGFLSTGFTMLALELCEIIVYGVMSDSYCR 240  
QY 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300  
DB 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300  
QY 301 TE 302  
DB 301 TE 302

## RESULT 3

US-09-895-298-66  
Sequence 66, Application US/09895298  
Publication No. US20030078405A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 47 Human Secreted Proteins  
FILE REFERENCE: P2035P1  
CURRENT APPLICATION NUMBER: US/09/895,298  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: 09/591,16  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: PCT/US99/29950  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: 60/113,006  
PRIOR FILING DATE: 1998-12-18  
PRIOR APPLICATION NUMBER: 60/112,809  
PRIOR FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 231  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 66  
LENGTH: 302  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (237)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-895-298-66

Query Match 99.8%; Score 1622; DB 10; Length 302;  
Best Local Similarity 99.7%; Pred. No. 6.2e-160; Indels 0; Gaps 0;  
Matches 301; Conservative 1; Mismatches 1;

QY 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVPGPLHFSGYG 60  
DB 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVPGPLHFSGYG 60  
QY 61 SVDPGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAFTVGFADVGQRST 120  
DB 61 SVDPGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAFTVGFADVGQRST 120  
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLQLTRMYPGLQ 180  
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLQLTRMYPGLQ 180



QY 181 VYFTEBMAVCDQIFODETGKRRSGSFLSTGCTMTLALCEBETIVYGVMSDYCR 240  
DB 181 VYFTEBMAVCDQIFODETGKRRSGSFLSTGCTMTLALCEBETIVYGVMSDYCR 240  
QY 241 EKSHPSVPHYFEKGRLECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSMR 300  
DB 241 EKSHPSVPHYFEKGRLECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSMR 300  
QY 301 TE 302  
DB 301 TE 302

RESULT 4  
US-09-895-298-120  
/ Sequence 120, Application US/09895298  
/ Publication No. US2003078405A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Rosen et al.  
/ TITLE OF INVENTION: 47 Human Secreted Proteins  
/ FILE REFERENCE: P2035P1  
/ CURRENT APPLICATION NUMBER: US/09/895,298  
/ PRIOR FILING DATE: 2001-07-02  
/ PRIOR APPLICATION NUMBER: 09/591,16  
/ PRIOR FILING DATE: 2000-06-09  
/ PRIOR APPLICATION NUMBER: PCT/US99/29950  
/ PRIOR FILING DATE: 1999-12-16  
/ PRIOR APPLICATION NUMBER: 60/113,006  
/ PRIOR FILING DATE: 1998-12-18  
/ PRIOR APPLICATION NUMBER: 60/112,809  
/ PRIOR FILING DATE: 1998-12-17  
/ NUMBER OF SEQ ID NOS: 231  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 120  
/ LENGTH: 345  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: SITE  
/ LOCATION: (280)  
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-895-298-120

Query Match 99.8%; Score 1622; DB 10; Length 345;  
Best Local Similarity 99.7%; Pred. No. 7.5e-160;  
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIILGVSVAAYIILCCWAGPLICLATGCLDHHPTGSRPTVPGPLHPSGYS 60  
DB 44 MKAPGRVLIIILGVSVAAYIILCCWAGPLICLATGCLDHHPTGSRPTVPGPLHPSGYS 103  
QY 61 SVDDGKPLVPEPCRSQAVSSSGQMLGSGAETIDSAECVFRMNOATVGFADVGQRST 120  
DB 104 SVDDGKPLVPEPCRSQAVSSSGQMLGSGAETIDSAECVFRMNOATVGFADVGQRST 163  
QY 121 LRVVSHSVPLLRNLSHYQKARDTLVYMWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180  
DB 164 LRVVSHSVPLLRNLSHYQKARDTLVYMWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 223  
QY 181 VYFTEBMAVCDQIFODETGKRRSGSFLSTGCTMTLALCEBETIVYGVMSDYCR 240  
DB 224 VYFTEBMAVCDQIFODETGKRRSGSFLSTGCTMTLALCEBETIVYGVMSDYCR 283  
QY 241 EKSHPSVPHYFEKGRLECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSMR 300  
DB 284 EKSHPSVPHYFEKGRLECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSMR 343  
QY 301 TE 302  
DB 344 TE 345

RESULT 5

US-10-262-511-184  
/ Sequence 184, Application US/10262511  
/ Publication No. US20040038223A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Smithson, Glenna  
/ APPLICANT: Millet, Isabelle  
/ APPLICANT: Peyman, John A.  
/ APPLICANT: Kekuda, Ramesh  
/ APPLICANT: Ju, Jingfang  
/ APPLICANT: Li, Li  
/ APPLICANT: Guo, Xiaojia (Sasha)  
/ APPLICANT: Patturajan, Meera  
/ APPLICANT: Spytek, Kimberly A.  
/ APPLICANT: Edinger, Shlomit R.  
/ APPLICANT: Ellerman, Karen  
/ APPLICANT: Malyankar, Uriel M.  
/ APPLICANT: Ort, Tatiana  
/ APPLICANT: Gorman, Linda  
/ APPLICANT: Zethusen, Bryan D.  
/ APPLICANT: Anderson, David W.  
/ APPLICANT: Zhong, Mei  
/ APPLICANT: Catterton, Elina  
/ APPLICANT: Miller, Charles E.  
/ APPLICANT: Rastelli, Luca  
/ APPLICANT: Stone, David J.  
/ APPLICANT: Pena, Carol E. A.  
/ APPLICANT: Shenoy, Suresh G.  
/ APPLICANT: Shinkets, Richard A.  
/ APPLICANT: Rothenberg, Mark E.  
/ APPLICANT: Leach, Martin D.  
/ APPLICANT: Agee, Michele L.  
/ APPLICANT: Berghe, Constance  
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
/ FILE REFERENCE: 21402-462C  
/ CURRENT APPLICATION NUMBER: US/10/262,511  
/ PRIOR FILING DATE: 2003-05-28  
/ PRIOR APPLICATION NUMBER: 60/326,483  
/ PRIOR FILING DATE: 2001-10-02  
/ PRIOR APPLICATION NUMBER: 60/373,815  
/ PRIOR FILING DATE: 2002-04-19  
/ PRIOR APPLICATION NUMBER: 60/327,917  
/ PRIOR FILING DATE: 2001-10-09  
/ PRIOR APPLICATION NUMBER: 60/381,642  
/ PRIOR FILING DATE: 2002-05-17  
/ PRIOR APPLICATION NUMBER: 60/328,029  
/ PRIOR FILING DATE: 2002-10-09  
/ PRIOR APPLICATION NUMBER: 60/381,038  
/ PRIOR FILING DATE: 2002-05-16  
/ PRIOR APPLICATION NUMBER: 60/328,056  
/ PRIOR FILING DATE: 2001-10-09  
/ PRIOR APPLICATION NUMBER: 60/373,260  
/ PRIOR FILING DATE: 2002-04-17  
/ PRIOR APPLICATION NUMBER: 60/373,826  
/ PRIOR FILING DATE: 2002-04-19  
/ PRIOR APPLICATION NUMBER: 60/327,435  
/ PRIOR FILING DATE: 2001-10-05  
/ Remaining Prior Application data removed - See File Wrapper or PALM.  
/ NUMBER OF SEQ ID NOS: 439  
/ SOFTWARE: Curaseqlist version 0.1  
/ SEQ ID NO 184  
/ LENGTH: 218  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-10-262-511-184

Query Match 71.8%; Score 1167; DB 15; Length 218;  
Best Local Similarity 100.0%; Pred. No. 9.2e-113;  
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 MLGSGGAEIIDSACVFRMNOATVGFADVGQRSTLRVVSHTSVPLLRNLSHYQKAR 144  
DB 1 MLGSGGAEIIDSACVFRMNOATVGFADVGQRSTLRVVSHTSVPLLRNLSHYQKAR 60

QY 145 DTLVWVGGRHMDRVLGRTYRTLLQTRMYPGLOVYTFTEEMAAVCDOIFODETGKNR 204  
 |||||  
 DB 61 DTLVWVGGRHMDRVLGRTYRTLLQTRMYPGLOVYTFTEEMAAVCDOIFODETGKNR 120  
 |||||  
 QY 205 RQSGSFLSTGFTMTLALCEIEIVYGVVSDSYCEKSHSPVHYHFEKGRIDECOMTL 264  
 |||||  
 DB 121 RQSGSFLSTGFTMTLALCEIEIVYGVVSDSYCEKSHSPVHYHFEKGRIDECOMTL 180  
 |||||  
 QY 265 AHEQAPSAHRFTTEKAVFSRMAKKRPVFAHPSWRT 302  
 |||||  
 DB 181 AHEQAPSAHRFTTEKAVFSRMAKKRPVFAHPSWRT 218  
 |||||

## RESULT 6

US-10-844-874-20  
 ; Sequence 20, Application US/10844874  
 ; Publication No. US20040204381A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moskal, Joseph  
 ; APPLICANT: Yamamoto, Hirotsuka  
 ; TITLE OF INVENTION: Detection and Treatment of Glyco-Enzyme-Related Disease  
 ; FILE REFERENCE: 97-186-E  
 ; CURRENT APPLICATION NUMBER: US/10/844,874  
 ; CURRENT FILING DATE: 2004-05-13  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 20  
 ; LENGTH: 305  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: rat alpha 2, 6-ST  
 ; US-10-844-874-20

Query Match 44.7%; Score 726.5; DB 17; Length 305;  
 Best Local Similarity 48.7%; Pred. No. 1,1e-66;  
 Matches 146; Conservative 46; Mismatches 85; Indels 23; Gaps 5;  
 QY 14 SVFSAVYILLCWAGLPLCLATCLDHHFPT-----GSRPTVPGLHFS-----GY 59  
 |||||  
 DB 12 AVSTALCIL-----LAMRLANDVTFPLLNCFQPKTKMLPLSTLRQLPQTHYGY 64  
 |||||  
 QY 60 SSVPDGKPLVPEPCRSQAVSSSGOMLGSLGAEIDSACVFRMNQAPTYGFEDVQGRS 119  
 |||||  
 DB 65 INVRTQEBL-QLNCHCAVAVNSGQVGVQKVEBIDRASCTWRMNNAPTKGFEEVDGYMT 123  
 |||||  
 QY 120 TLRAVSHSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQTRMYPG 179  
 |||||  
 DB 124 WRAVSHSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQTRMYPG 183  
 |||||  
 QY 180 QVYTFTEEMAAVCDOIFODETGKNRROSGSFLSTGFTMTLALCEIEIVYGVVSDSYC 239  
 |||||  
 DB 184 QIVYTFTEEMAAVCDOIFODETGKNRROSGSFLSTGFTMTLALCEIEIVYGVVSDSYC 243  
 |||||  
 QY 240 REKSHSPVHYHFEKGRIDECOMTLAHEQAPSAHRFTTEKAVFSRMAKKRPVFAHPSW 299  
 |||||  
 DB 244 TTEGYRKVPYHYEQQGK-DECNEYLHSHADYGGHRFTTEKAVFAKMAKKRIVFTHPNW 302  
 |||||

## RESULT 7

US-09-833-245-1914  
 ; Sequence 1914, Application US/09833245  
 ; Publication No. US20040010134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PFS46PCT  
 ; CURRENT APPLICATION NUMBER: US/09/833,245  
 ; CURRENT FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229, 358  
 ; PRIOR FILING DATE: 2000-04-12

;; PRIOR APPLICATION NUMBER: 60/256, 931  
 ;; PRIOR FILING DATE: 2000-12-21  
 ;; PRIOR APPLICATION NUMBER: 60/199, 384  
 ;; PRIOR FILING DATE: 2000-04-25  
 ;; NUMBER OF SEQ ID NOS: 2267  
 ;; SOFTWARE: PatentIn Ver. 2.1  
 ;; SEQ ID NO 1914  
 ;; LENGTH: 305  
 ;; TYPE: PRT  
 ;; ORGANISM: Homo sapiens  
 ; US-09-833-245-1914

Query Match 44.1%; Score 716.5; DB 11; Length 305;  
 Best Local Similarity 52.7%; Pred. No. 1,2e-65;  
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;

QY 55 HFSGYSSVPDGKPLVPEPCRSQAVSSSGOMLGSLGAEIDSACVFRMNQAPTYGFED 114  
 |||||  
 DB 61 HY-GYINVTQEBPLQD-CDLCAIVNSGQVGVQKVEBIDRASCTWRMNNAPTKGYEED 118  
 |||||  
 QY 115 VGORSTLRVSHSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQTR 174  
 |||||  
 DB 119 VGRMTLRVSHSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQTR 178  
 |||||  
 QY 175 MYPGLOVYTFTEEMAAVCDOIFODETGKNRROSGSFLSTGFTMTLALCEIEIVYGVV 234  
 |||||  
 DB 179 IYFNAQIVYTFTEEMAAVCDOIFODETGKNRROSGSFLSTGFTMTLALCEIEIVYGVV 238  
 |||||  
 QY 235 SDSYCEKSHSPVHYHFEKGRIDECOMTLAHEQAPSAHRFTTEKAVFSRMAKKRPVYF 294  
 |||||  
 DB 239 NDTYCTEGYRKVPYHYEQQGK-DECNEYLHSHADYGGHRFTTEKAVFAKMAKKRILF 297  
 |||||  
 QY 295 AHPWSW 299  
 |||||  
 DB 298 THPNW 302  
 |||||

## RESULT 8

US-09-833-245-1915  
 ; Sequence 1915, Application US/09833245  
 ; Publication No. US20040010134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PFS46PCT  
 ; CURRENT APPLICATION NUMBER: US/09/833,245  
 ; CURRENT FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229, 358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256, 931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199, 384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1915  
 ; LENGTH: 305  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-833-245-1915

Query Match 44.1%; Score 716.5; DB 11; Length 305;  
 Best Local Similarity 52.7%; Pred. No. 1,2e-65;  
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;

QY 55 HFSGYSSVPDGKPLVPEPCRSQAVSSSGOMLGSLGAEIDSACVFRMNQAPTYGFED 114  
 |||||  
 DB 61 HY-GYINVTQEBPLQD-CDLCAIVNSGQVGVQKVEBIDRASCTWRMNNAPTKGYEED 118  
 |||||  
 QY 115 VGORSTLRVSHSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQTR 174  
 |||||  
 DB 119 VGRMTLRVSHSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQTR 178  
 |||||

QY 175 MYBGLQVYTFTEEMMAVCDOIFODETGKRNROSGSFISTGFMITLALCEBIIVYGVW 234  
DB 179 IYRNAQIYVTTTEKMSYCDVFKETGKDRVQSGSYLSTGFTFLAMDACYGHIHYGMI 238  
QY 235 SDSYCREKSHSPVYHYFEKGRUDECOMYLAHQAPRSARHFTTEKAVFSRMAKKRPIVF 294  
DB 239 NDTYCKTEGKRVKPYHYHYEGQR-DECDYFLHBAHPGGRHFTTEKKVFAKMAKKRIIF 297  
QY 295 AHPSW 299  
DB 298 THPNW 302

## RESULT 9

US-10-188-186-22  
; Sequence 22, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; PRIOR FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 22  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-186-22

Query Match 43.9%; Score 714.5; DB 15; Length 305;  
Best Local Similarity 52.2%; Pred. No. 1.9e-65;  
Matches 128; Conservative 49; Mismatches 65; Indels 3; Gaps 3;  
QY 55 HFSGYSSVDPDKPLVREBPCRSACAVVSSGQMLSGLAEDISABCVPRNQAFTVGEAD 114  
DB 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQWGVQKVGNEIDRSSCIWRMNAAPTGYEED 118  
QY 115 VQGRSTLRVVSHTSVPLLLRNYSHYFOKADTLVMWGGQRANDRVLGRTYRTLQLTR 174  
DB 119 VGMRTIRVVSHTSVPLLLRNPDYFFKEANTTIYVIGPFRNRKQDNGIYNNMLKKTVG 178  
QY 175 MYBGLQVYTFTEEMMAVCDOIFODETGKRNROSGSFISTGFMITLALCEBIIVYGVW 234  
DB 179 IYRNAQIYVTTTEKMSYCDVFKETGKDRVQSGSYLSTGFTFLAMDACYGHIHYGMI 238  
QY 235 SDSYCREKSHSPVYHYFEKGRUDECOMYLAHQAPRSARHFTTEKAVFSRMAKKRPIVF 294  
DB 239 NDTYCKTEGKRVKPYHYHYEGQR-DECDYFLHBAHPGGRHFTTEKKVFAKMAKKRIIF 297  
QY 295 AHPSW 299

DB 298 THPNW 302

## RESULT 10

US-10-104-047-2350  
; Sequence 24, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2350  
; LENGTH: 305  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-2350

Query Match 43.9%; Score 713.5; DB 14; Length 305;  
Best Local Similarity 52.2%; Pred. No. 2.4e-65;  
Matches 128; Conservative 49; Mismatches 65; Indels 3; Gaps 3;

QY 55 HFSGYSSVDPDKPLVREBPCRSACAVVSSGQMLSGLAEDISABCVPRNQAFTVGEAD 114  
DB 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQWGVQKVGNEIDRSSCIWRMNAAPTGYEED 118  
QY 115 VQGRSTLRVVSHTSVPLLLRNYSHYFOKADTLVMWGGQRANDRVLGRTYRTLQLTR 174  
DB 119 VGMRTIRVVSHTSVPLLLRNPDYFFKEANTTIYVIGPFRNRKQDNGIYNNMLKKTVG 178  
QY 175 MYBGLQVYTFTEEMMAVCDOIFODETGKRNROSGSFISTGFMITLALCEBIIVYGVW 234  
DB 179 IYRNAQIYVTTTEKMSYCDVFKETGKDRVQSGSYLSTGFTFLAMDACYGHIHYGMI 238  
QY 235 SDSYCREKSHSPVYHYFEKGRUDECOMYLAHQAPRSARHFTTEKAVFSRMAKKRPIVF 294  
DB 239 NDTYCKTEGKRVKPYHYHYEGQR-DECDYFLHBAHPGGRHFTTEKKVFAKMAKKRIIF 297  
QY 295 AHPSW 299  
DB 298 THPNW 302  
RESULT 11  
US-10-188-186-24  
; Sequence 24, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; PRIOR FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11

```

PRIORITY APPLICATION NUMBER: 60/305262
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/373881
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/305673
PRIOR FILING DATE: 2001-07-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 368
SOFTWARE: Custom
SEQ ID NO 24
LENGTH: 318
TYPE: PROT
ORGANISM: Homo sapiens
US-10-188-186-24

Query Match      42.7%; Score 694; DB 15; Length 318;
Best Local Similarity 49.6%; Pred. No. 2,86-63;
Matches 128; Conservative 48; Mismatches 66; Indels 16; Gaps 4

55 HFSGYSVPDDGKPLTVERPCRSQCAVYSSSGOMLGSGLGAEIDSACVFRMNOAPTVGFEAD 114
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
61 HY-GYINKKTOGEPLOLD-CDLCAIVNSGQWGVKQKNEIDRSSCIWRNNAPATGYEED 118

115 VGORSTLAVSHTSYPLILIRYNSHIFQKARDPLNWWGGGRMDVIGGRTYRTLLQLR 174
119 VGRMTMIRVSTSVPLILKNDPYFKEKNTITTYIYWGPFRRNRKQGNQIVYNNLKKTVG 178

175 MYPGLOVYTFETRMMAVACDQIFQDETGNKR-----RSGSFLSTGWFPTMILA 221
179 IYPAQIYVTTTERKMSYCDGCVFKETGKDSVLTLSKSSFOIVGSGVYLSGWFPTILA 238

222 LELCEIIVYGVNSDYSYCKEKSPSYPHYIYFEKGRILDEQMTLANHQAARSHRFTTEKA 281
229 MDACGIIHYGGIMINITYCTKEGKYRVPHYVYQGR-DECEYFLHEHAPYGGHRTTEKK 297

QY 282 VFSRAKKRPPIEFAHPSW 299
DB 298 VFAKMAKKRITLFTHEPW 315

RESULT 12
US-10-188-186-20
Sequence 20, Application US//10188186
Publication No. US20040029789A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-397C
CURRENT APPLICATION NUMBER: US/10/188,186
CURRENT FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 60/303046
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/360814
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/303828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 60/323380
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/361133
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/340416
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/344502
PRIOR FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 60/305262
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/373881
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/305673
PRIOR FILING DATE: 2001-07-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 368
SOFTWARE: Custom

```

```

; SEQ ID NO 20
; LENGTH: 305
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-188-186-20

Query Match      41.7%; Score 677.5; DB 15; Length 305;
Best Local Similarity 49.4%; Pred. No. 1,4e-61;
Matches 121; Conservative 49; Mismatches 72; Indels 3; Gaps 3;

QY HFSGYSVPDGPCLVRRPCSCAVVSSSGGMTSGLAETSDASAECPFRNQAFTVGFEAD 114
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB HY-GYINVKOEPLQLD-CDLCAIVNSGOMGVKQKVENEIDRSSCIWRNNNAFTTKVEED 118
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY VGORSTLRVSVSHTSVPLLRLRYSHYFGKARDLIYMWGGGRHMDRVLGRTYTTLQLTR 114
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB VGRMTMIRVSVHSHTSVPLLKNPVDYEFKAENTTIYIVGWPPRNMRKDNGIYVMNLKKTGV 178
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY MYPLGLQVTFPERMMAYCDQIFODETGNKRQSGSFISTGMFMITLALCEEIVYGVW 234
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB IYPNAQIVLTTEKMSSCYDGVFKEKETGDSIEHAVLIDQGFFITLMADACIGIHVGM 238
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY SDSVCREKSHPSVPVHYEFEGKRLDECOMYLAEQAPRSARHFITEKAFVSMAKKRPDV 294
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB NDLYKTEGTGYKVRPHYHEQGR-DEQDEYFLHEHPA PGHRPFITEKVPFAKMAKKRIIF 297
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY AHPWS 299
   |||
DB 298 THPNW 302

RESULT 13
US-10-097-340-189
; Sequence 189, Application US/10097340
; Publication No. US20030087250a1
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNANVARAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shudhangt KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. EAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification
TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0

```

SEQ ID NO 189  
LENGTH: 336  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-097-340-189

Query Match 39.3%; Score 639.5; DB 14; Length 336;  
Best Local Similarity 47.9%; Pred. No. 1.4e-57;  
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPYTP-GPLHFGSYSDVDGKPLVREPCSCAVVSSGQMLGSLGAIDSAECVFR 102  
DB 64 TQGRPGVAPRPLDGLGVADHKPL-KMCRDCAVTSSGHLHSGSGSIDOTECVIR 122  
QY 103 MNOAPTVGEADYGQSTLHVSHTSVPLLRNYSHYFQKARDTLVNWGGGRHMDVILG 162  
DB 123 MNDAPRTGYGRDVGNRTSLVIAHSSIQRLRNHDLNVSGQTVLFWGSSYMRDCK 182  
QY 163 GRYYRTLQLTMYPGLOVYTFERRMAYCDQIFODETGKRRSGSFLSTGWTMLAL 222  
DB 183 GQVYNNHLHLSQVLPRLKAMITRHKMLQDFDEKQETGDKRKISNTMLSTGWTMLAL 242  
QY 223 ELCEEIVYGVMSDSYCREKSHSPVPHYFEKGRLDCEQMYLAHQAPR-SAHRFITEKA 281  
DB 243 ELCDRINVGMPDPFCRDENHPSVPHYHYEPGPDCTMYLSHERGKSGSHRFITEKR 302  
QY 282 VFSRMAKRPIVFAHPSWTE 302  
DB 303 VFKMWARTFNIHFQPDWKPE 323

RESULT 14  
US-10-094-749-2441  
Sequence 2441, Application US/10094749  
Publication No. US20030219741A1

GENERAL INFORMATION:

APPLICANT: ISOGAI, TAKAO  
APPLICANT: SUGIYAMA, TOMOYASU  
APPLICANT: OTSUKI, TETSUJI  
APPLICANT: WAKAMATSU, AI  
APPLICANT: SATO, HIROYUKI  
APPLICANT: ISHII, SHIZUKO  
APPLICANT: YAMAMOTO, JUN-ICHI  
APPLICANT: ISONO, YUUKO  
APPLICANT: HIO, YURI  
APPLICANT: OTSUKA, KAORU  
APPLICANT: NAGAI, KEIICHI  
APPLICANT: IRIE, RYOTARO  
APPLICANT: TAMECHIKA, ICHIRO  
APPLICANT: SEKI, NAOHITO  
APPLICANT: YOSHIKAWA, TSUTOMU  
APPLICANT: OTSUKA, MOTOYUKI  
APPLICANT: NAGAHARI, KENJI  
APPLICANT: MASUHO, YASUHIKO  
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
FILE REFERENCE: 084335/0160  
CURRENT APPLICATION NUMBER: US/10/094,749  
PRIOR FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: 60/350,435  
PRIOR FILING DATE: 2002-01-24  
PRIOR APPLICATION NUMBER: JP 2001-328381  
PRIOR FILING DATE: 2001-09-14  
NUMBER OF SEQ ID NOS: 3381  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2441  
LENGTH: 336  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-094-749-2441

Query Match 39.3%; Score 639.5; DB 14; Length 336;  
Best Local Similarity 47.9%; Pred. No. 1.4e-57;  
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPYTP-GPLHFGSYSDVDGKPLVREPCSCAVVSSGQMLGSLGAIDSAECVFR 102  
DB 64 TQGRPGVAPRPLDGLGVADHKPL-KMCRDCAVTSSGHLHSGSGSIDOTECVIR 122  
QY 103 MNOAPTVGEADYGQSTLHVSHTSVPLLRNYSHYFQKARDTLVNWGGGRHMDVILG 162  
DB 123 MNDAPRTGYGRDVGNRTSLVIAHSSIQRLRNHDLNVSGQTVLFWGSSYMRDCK 182  
QY 163 GRYYRTLQLTMYPGLOVYTFERRMAYCDQIFODETGKRRSGSFLSTGWTMLAL 222  
DB 183 GQVYNNHLHLSQVLPRLKAMITRHKMLQDFDEKQETGDKRKISNTMLSTGWTMLAL 242  
QY 223 ELCEEIVYGVMSDSYCREKSHSPVPHYFEKGRLDCEQMYLAHQAPR-SAHRFITEKA 281  
DB 243 ELCDRINVGMPDPFCRDENHPSVPHYHYEPGPDCTMYLSHERGKSGSHRFITEKR 302  
QY 282 VFSRMAKRPIVFAHPSWTE 302  
DB 303 VFKMWARTFNIHFQPDWKPE 323

RESULT 15  
US-10-288-252-11  
Sequence 11, Application US/10288252  
Publication No. US20030143686A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: LAL, Preeti G.  
APPLICANT: TANG, Y. Tom  
APPLICANT: YUE, Henry  
APPLICANT: BURFORD, Neil  
APPLICANT: GANDHI, Ameeta R.  
APPLICANT: WARREN, Bridget A.  
APPLICANT: YAO, Monique G.  
APPLICANT: TRIBOUTLEY, Catherine M.  
APPLICANT: BAUGHN, Mariah R.  
APPLICANT: LEE, Ernestine A.  
APPLICANT: HAFALIA, April J.A.  
APPLICANT: LU, Yan  
APPLICANT: GRIFFIN, Jennifer A.  
APPLICANT: SANJUNWALA, Madhu S.  
APPLICANT: DING, Li  
TITLE OF INVENTION: TRANSFERASES  
FILE REFERENCE: PI-0241 USA  
CURRENT APPLICATION NUMBER: US/10/288,252  
PRIOR FILING DATE: 2002-11-04  
PRIOR APPLICATION NUMBER: PCT US01/30424  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: US 60/252,819  
PRIOR FILING DATE: 2000-11-21  
PRIOR APPLICATION NUMBER: US 60/249,639  
PRIOR FILING DATE: 2000-11-16  
PRIOR APPLICATION NUMBER: US 60/247,931  
PRIOR FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: US 60/246,001  
PRIOR FILING DATE: 2000-11-03  
PRIOR APPLICATION NUMBER: US 60/244,025  
PRIOR FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: US 60/238,481  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: US 60/236,523  
PRIOR FILING DATE: 2000-09-29  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PERL Program  
SEQ ID NO 11  
LENGTH: 336  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030143686A1 2778782CD1  
US-10-288-252-11

Query Match 39.1%; Score 636.5; DB 14; Length 336;  
 Best Local Similarity 47.5%; Pred. No. 2,9e-57;  
 Matches 124; Conservative 45; Mismatches 89; Indels 3; Gaps 3;

QY	44	TGSRPTVP-GPLHPSGYSSVPDGGKPLYREPCRCACAVSSGOMLGSLGAELIDSAECVFR	102
DB	64	TQORPGVPAGPRPLDGYLGADHKPL-KMHCRCALVTSGHLLHSRQSQIDQTECVIR	122
QY	103	MNQAPTVGFADYGGORETLTVSHTSVPLLRNYSHYFCKARDPLVVMWGQGRHMDRVLG	162
DB	123	MNDAPTKGYGRDVGNRTSLKRVIAHSSIQRIKRNHDLINVSOGTVFIWGPSSYMRDGG	182
QY	163	GRTYRTLLQLTRMYPGLQVYTFTERMMAYCDQIFODETGKRRRQSGSFLSTGFTMIAL	222
DB	183	GQYYNNLHLISQVLPRLKAFMTIRHKTLQFDELFKQETGKORKISNTWLSTGFTMTIAL	242
QY	223	ELCEELVYGMVSDSYCREKSHPSVPYHYFEKGRLDCCOMYLAHQAPR-SAHREITEKA	281
DB	243	ELCDRIINVGVVPPDFCRDPNHPSPYHYDPFGPDCTMYLSHERGKSGSHRFLTEKR	302
QY	282	VFSRWAKKRPIVFAHPSMRTE	302
DB	303	VFKNMARTFNHIFQPDMPKE	323

Search completed: December 21, 2004, 22:50:58  
 Job time : 79 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 21, 2004, 22:44:46 ; Search time 105 Seconds  
(without alignments)  
2044.364 Million cell updates/sec

Title: US-10-019-114A-7  
Perfect score: 1626  
Sequence: 1 MKRAGRLLVILLCVVFSAV.....FSRMAKRPIVFAHPSWRT 302

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODE=frame+p2n.model -DEV=x1h  
-Q=/cgn2\_1/USPTO.spool/US10019114/runat\_21122004\_100141\_3443/app\_query.fasta\_1.455  
-DB=Issued Patents NA -OPMT=faetap -SUFFIX=rml -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdt  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTMT=p2n -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10019114@cgn\_1.1\_69@runat\_21122004\_100141\_3443 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -MNG SCORES=0 -NAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1626	100.0	1556	4	US-09-599-360B-25
2	639.5	39.3	2056	3	US-09-334-601-12
3	544.5	33.5	2697	4	US-09-620-312D-401
4	494	30.4	546	4	US-09-621-976-55
5	196.5	12.1	2026	1	US-08-503-133A-1
6	196.5	12.1	2026	2	US-08-576-775A-1
7	196.5	12.1	2026	2	US-08-972-498-1
8	196.5	12.1	2026	3	US-08-899-545-1
9	194.5	12.0	1020	2	US-08-446-875-15
10	192.5	11.8	1218	2	US-08-446-875-1
11	192.5	11.8	1218	2	US-08-102-385G-1
12	190.5	11.7	1048	1	US-08-626-994A-4

13	190.5	11.7	1048	3	US-08-957-742-4	Sequence 4, Appl
14	190.5	11.7	1660	1	US-08-626-994A-2	Sequence 2, Appl
15	190.5	11.7	1660	3	US-08-957-742-2	Sequence 2, Appl
16	188.5	11.6	1292	1	US-08-503-133A-7	Sequence 7, Appl
17	188.5	11.6	1292	2	US-08-576-775A-7	Sequence 7, Appl
18	188.5	11.6	1292	2	US-08-972-498-7	Sequence 7, Appl
19	188.5	11.6	1292	3	US-08-899-545-7	Sequence 7, Appl
20	188.5	11.6	1682	1	US-08-503-133A-8	Sequence 8, Appl
21	188.5	11.6	1682	2	US-08-576-775A-8	Sequence 8, Appl
22	188.5	11.6	1682	2	US-08-972-498-8	Sequence 8, Appl
23	188.5	11.6	1682	3	US-08-899-545-8	Sequence 8, Appl
24	179	11.0	1188	2	US-08-446-875-9	Sequence 9, Appl
25	179	11.0	1188	2	US-08-102-385G-9	Sequence 9, Appl
26	178	10.9	1206	2	US-08-666-367B-2	Sequence 2, Appl
27	178	10.9	1206	3	US-09-143-438-2	Sequence 2, Appl
28	178	10.9	2671	2	US-08-666-367B-1	Sequence 2, Appl
29	178	10.9	2671	3	US-09-143-438-1	Sequence 1, Appl
30	176	10.8	1206	2	US-08-446-875-3	Sequence 3, Appl
31	176	10.8	1206	2	US-08-102-385G-3	Sequence 3, Appl
32	174.5	10.7	1632	2	US-09-334-601-9	Sequence 9, Appl
33	172	10.6	2117	4	US-08-361-304A-1	Sequence 1, Appl
34	171	10.5	1146	2	US-08-666-367B-4	Sequence 4, Appl
35	171	10.5	1146	3	US-09-143-438-4	Sequence 4, Appl
36	168	10.3	1246	1	US-08-446-777-3	Sequence 3, Appl
37	168	10.3	1540	4	US-09-023-655-1250	Sequence 1250, Ap
38	168	10.3	2304	1	US-08-446-777-5	Sequence 5, Appl
39	168	10.3	2304	1	US-08-446-777-7	Sequence 7, Appl
40	167	10.3	2528	4	US-10-140-002-461	Sequence 461, App
41	150.5	9.3	1158	2	US-08-446-875-11	Sequence 11, Appl
42	150.5	9.3	1158	2	US-08-102-385G-11	Sequence 11, Appl
43	150.5	9.3	1766	1	US-07-991-587A-6	Sequence 6, Appl
44	150.5	9.3	1766	1	US-08-309-985-6	Sequence 6, Appl
45	149	9.2	2121	4	US-09-425-488-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-09-599-360B-25  
; Sequence 25, Application US/09599360B  
; Patent No. 6548633  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Bouguetier, L.  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: GENSET.050CP3  
; CURRENT APPLICATION NUMBER: US/09/599,360B  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/113,686  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/141,032  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 09/469,099  
; PRIOR FILING DATE: 1999-12-21  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: Patent.pm  
; SEQ ID NO 25  
; LENGTH: 1556  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 261..1166  
; NAME/KEY: sig\_peptide  
; LOCATION: 261..314  
; OTHER INFORMATION: Von Heijne matrix  
; OTHER INFORMATION: score 8.80  
; NAME/KEY: polyA\_site  
; LOCATION: 1524..1556  
; US-09-599-360B-25

## Alignment Scores:

Pred. No.: 3,61e-179 Length: 1556  
 Score: 1626.00 Matches: 302  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-019-114a-7 (1-302) x US-09-599-360B-25 (1-1556)

```

QY 1 MetLysAlaProGlyArgLeuValLeuIleLeuCySerSerValAlaPheSerAlaVal 20
Db ATGAAGGCTCGGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 320
QY 21 TyrlleLeuLeuCySerSTPAlaGlyLeuProLeuCySerLeuAlaThrCysLeuAspHis 40
Db TACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
QY 41 HisPheProThrGlySerArgProThrValProGlyProLeuHisPheSerGlyTyrSer 60
Db CACTTCCCAAGGCTCCAGGCCCACTGTGCGGAGCCCTTGCATTCACTGATATAGC 440
QY 61 SerValProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValSer 80
Db AGTGTCCAGATGGGAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
QY 81 SerSerGlyGlnMetLeuGlySerGlyLeuGlyValAlaGluIleAspSerAlaGluCysVal 100
Db AGTGTCCGCGCAAAATGCTGGGCTCAGGCTGGGCTGCTAGATCCAGATGCGAGTGGTG 560
QY 101 PheArgMetAsnGlnAlaProThrValGlyPheGlnAlaAspValGlyIleArgSerThr 120
Db TTCCGATGAAACAGAGGCCCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 620
QY 121 LeuArgValAlaSerHisThrSerValProLeuLeuLeuArgAntyrSerHisTyrPhe 140
Db CTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680
QY 141 GlnLysAlaArgAspThrLeuTyrMetValTyrGlyGlnIleArgHisSmetAspArgVal 160
Db CAGAAGGCCGGAAGCAGGCTCTCATGTGTGGGCGAAGGCGAAGGCGAAGGCGAGG 740
QY 161 LeuGlyValArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGln 180
Db CTGGGGGCGGCGACCTACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
QY 181 ValTyrThrPheThrGluArgMetLeuAlaTyrCysAspGlnIlePheGlnAspGluThr 200
Db GTGTACACCTTCAAGGAGCGCATGATGCTTACGACCAAGATCTTCCAGAGCAAGACG 860
QY 201 GlyLysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeu 220
Db GGCAAGAACCGGAGGAGTGGGCTCTCTTCCAGACCGGCTGCTTCCACATGATCTTC 920
QY 221 AlaLeuGlnLeuCySerGluGlnIleValIleValTyrGlyMetValSerAspSerTyrCysArg 240
Db GCCTGAGAGCTGTGTGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 980
QY 241 GlnLysSerHisProSerValProTyrHisTyrPheGlnLysGlyArgLeuAspGluCys 260
Db GAGAAGAGCCACCCCTCAGTGCCTTACCACTTATAGAAGGCGCGGTAGATGATGT 1040
QY 261 GlnMetTyrLeuAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluLys 280
Db GAGATGATCTGCGACAGCAGAGCGCCCGGAGCCGCCACCGCTTCACTACAGAGAG 1100
QY 281 AlaValPheSerArgTyrPalaLysLysArgProIleValPheAlaHisProSerTyrArg 300
Db GCGGCTCTTCTCCGCTGGGCGCAAGAGGCCCATCTGTGTCGCCCATCTGCTGAGG 1160
QY 301 ThrGlu 302
Db 1161 ACTGAG 1166

```

## RESULT 2

US-09-334-601-12  
 ; Sequence 12, Application US/09334601

; Patent No. 6280989  
 ; GENERAL INFORMATION:

; APPLICANT: Kapilnov, Dmitri

; TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES

; FILE REFERENCE: VCUIP-6

; CURRENT APPLICATION NUMBER: US/09/334,601

; CURRENT FILING DATE: 1999-06-17

; NUMBER OF SEQ ID NOS: 94

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 12

; LENGTH: 2056

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (181)..(1188)

US-09-334-601-12

US-10-019-114a-7 (1-302) x US-09-334-601-12 (1-2056)

Alignment Scores:

Pred. No.: 4.01e-64 Length: 2056  
 Score: 639.50 Matches: 125  
 Percent Similarity: 64.75% Conservative: 44  
 Best Local Similarity: 47.89% Mismatches: 89  
 Query Match: 39.33% Indels: 3  
 DB: 3 Gaps: 3

```

QY 44 ThrGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTyrSerSerVal 62
Db ACCAGAGGCGCCCGGGGTCCCGGCGGAGCGGCGCACTGACGATACCTCGGAGTG 429
QY 63 ProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValSerSerSer 82
Db GCGGACCAAGCCCTG---AAATGCACTGCAAGGAGCTGTCTGCTGCTGCTGCTG 486
QY 83 GlyLysMetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCysValPheArg 102
Db GGCATCTGTGCAAGTGGCAAGGCTCCAGATGACACAGAGTGTGTATCGCG 546
QY 103 MetAsnGlnAlaProThrValGlyPheGlnAlaAspValGlyIleArgSerThrLeuArg 122
Db ATGATGACCGCCCGCACCGGCTATGGCGTGAAGTGGCAATGCGCACACCTGAGG 606
QY 123 ValAlaSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPheGlnLys 142
Db GTATCGCGCATTCACATCCAGATCCGAGATCTTCCGACCGCATGATCTGCTGCAAGTG 666
QY 143 AlaArgAspThrLeuTyrMetValTyrGlyGlnIleArgHisMetAspArgValLeuGly 162
Db AGCAGAGGACCGGTTCATCTTCTGGGCGCCAGCAGTACATGCGGCGGAGCGGCAAG 726
QY 163 GlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyr 182
Db GGCCAGGCTTACAAACACTGCACTCTGAGCAGAGTGTGCGCCGCGGAGGCGCTTC 786
QY 183 ThrPheThrGluArgMetValTyrCysAspGlnIlePheGlnAspGluThrGlyLys 202
Db ATGATTAATCTGCCACAAGATGCTGCACTTGTGATGAGTCTTCAAGCAGAGATCTGCAAA 846
QY 203 AsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeu 222
Db GACAGAGAGATATCCAAACCTTGCTCAGACCTGGCTTTCATATGACATGCAATGCACTG 906
QY 223 GlnLeuCySerGluGlnIleValValTyrGlyMetValSerAspSerTyrCysArgGluLys 242
Db GAGCTGTGACAGATCAATGTTATGAGATGATGCGCCCGCAGACTTTCAGAGGATCCC 966

```



Db 359 CCGTACCGCGCACTGTCAACCTTACAGAGTGAAGCATCTACAGCGGCTATGTCCCAAT 418

Qy 62 LProaerpglyysProleuValaLarglupProCyaaRgsersCyAAlaValaIserse 82

Db 419 TCTCGGCAACAAGACATGCTCCCTCCG---TGCACACAGTGTGGATATTGTACAGAGCTC 475

Qy 82 rGlyImetleuglyIserglyLeuglyAlaGluIleapserAlaGluCyvalaPhear 102

Db 476 CAGCCACCTGCTGGGCAACCAAGCTGGGCGCTGAAGATCCAGCGGCTGAAGTATCAATCCG 535

Qy 102 gMeRanpnlAaProthrrValaGlyPheGluAaIaerValaGlyIlnaRgsertThreulr 122

Db 536 CATGAATATATGACACCCACCACTGGCTACTACACTATATGGGCAACAAGACCACTTACCG 595

Qy 122 gValaIserhIsthrSerValaProleuIleuIleuIleuIleuIleuIleuIleuIleu 142

Db 596 CCGTCTGGCGCCATTCCAGTGTCTTCCGCGCTGTGAAGAGCGCCCAAGAGTGTGTACACGG 655

Qy 142 sAlaIaRgaPthreulrYrMetValaTrpGlyGluIlnaRghIsmetAaRgaValaIleu 162

Db 656 GACCCCTGAACCCGHTTCATCTTCTGGGGGCGCCCGCAAGATGACAGAACCCCGAGG 715

Qy 162 yGlyaRghrYrAaThrrleuIleuGluIleuThrrRgmetyrTrpGlyLeuGluVala 182

Db 716 CAGC---CTCGTCCGTGTGATCCAGCGGCGCGCTGTGTCCCAACATGGAAGCAT 772

Qy 182 rThrrPheRghrGluAaRgmetyrMetAlaTrCyaaRghIlePheGluAaRghIuthrGly 202

Db 773 TGCCGCTCTCCCGCGCGCATGCGGCATTTGACACACTCTTCCGGGGTGAACGGGCA 832

Qy 202 sAaNaRgaRghIserglySerPheIleuSerThrGlyTrpPheThrMetIleuAla 222

Db 833 GGACAGAGAGAACTCATCTCGTGGTGAACAAGCTGGTTCATCCATGATGATGCGCGT 892

Qy 222 uGluIleuCyGluIleuValaValaTrpGlyMetAlaSerAspserTrpCyAaRghIu 242

Db 893 GGAGTGTGTGACCAAGCATGTCTATGACATGGTGTCCCGCCCACTACTGACGCAGG 952

Qy 242 sSerHis---ProserValaProYrHIsTrYrPheGluIleuIleuIleuIleuIleuIleu 261

Db 953 GCGCCGCTCCAGCGCATGCGCTTACCTACCTACAGACCCCAAGGGCGCGAGAAATGT 101

Qy 261 mMetYrleuAlaHIsGluGluAlaProArg---SerAlaHIsaRghIeThrGly 280

Db 1013 CACCTACATCCAGATGAGCAACAGTGGAGGCAACCAACCGCTTCATACCGAGAA 107

Qy 280 sAlaValaPheSerAgtPAlaIleuIleuIleuIleuIleuIleuIleuIleuIleu 299

Db 1073 AAGGTCCTTCATCGTGGCGCCAGCTGATGACATCCCTTCCACCCCTCTCTGG 1130

RESULT 4

US-09-621-976-55

/ Sequence 55, Application US/09621976

/ Patent No. 6639063

/ GENERAL INFORMATION:

/ APPLICANT: Dumas Milne Edwards, J.B.

/ APPLICANT: Jober, S.

/ APPLICANT: Giordano, J.Y.

/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.

/ FILE REFERENCE: GENSET.054PR2

/ CURRENT APPLICATION NUMBER: US/09/621.976

/ CURRENT FILING DATE: 2000-07-21

/ NUMBER OF SEQ ID NOS: 19335

/ SOFTWARE: Patent.pm

/ SEQ ID NO 55

/ LENGTH: 546

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: 260..544

/ NAME/KEY: sig\_peptide



## RESULT 6

US-08-576-775A-1

Sequence 1, Application US/08576775A

Patent No. 5849904

GENERAL INFORMATION:

APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;

APPLICANT: Nakayama, Jun; Eckhardt, Matthias

TITLE OF INVENTION: Isolated Polysialyl Transferases,

TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felte &amp; Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/576, 775A

FILING DATE: 21-December-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/503,133

FILING DATE: 17-July-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP94/04289

FILING DATE: 22-December-1994

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5849904man D.

REGISTRATION NUMBER: 30,946

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2026 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 301..1377

US-08-576-775A-1

## Alignment Scores:

Pred. No.: 1,47e-12 Length: 2026

Score: 196.50 Matches: 71

Percent Similarity: 45.93% Conservative: 42

Best Local Similarity: 28.86% Mismatches: 99

Query Match: 12.08% Indels: 35

DB: 2 Gaps: 8

US-10-019-114A-7 (1-302) x US-08-576-775A-1 (1-2026)

QY 59 TyrSerSerValProAspGlyLysProLeuValArgLysIleProCysArgSerCysAlaVal 78

DB 673 CACAGCCTCGCTGCTGAAGTTTCAACCAATGAAAAACCGCAGCTTAAAGACTGCTGCTT 732

QY 79 ValSerSerSerGlyMetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGlu 98

DB 733 GTTGAAACTCTGCGATTCTAGACAGTGTGCAAGGAGATTGACAGTCACAT 792

QY 99 CysValPheArgMetLeuGlnIleProThrValGlyPheGlnIleAspValGlyGlnArg 118

DB 793 TTTGTAATCAGGTGCAATCTAGCTCTGCTGCTGAGATTGCTGCCGAGATGCGCACTAA 852

QY 119 SerThrLeuArgValValSerHisThrSerValProLeuLeu-----LeuArgAsn 135

DB 853 TCAGATTTTATTACCATGAACCCATCAGTTGTGCGAGAGACATTTGAGGCTTTGGAT 912

QY 136 TyrSer-----HisThrPheGlnIleValAlaArgPheLeu----- 147

DB 913 GAGAGTCACAGACGAAATTTGCTATGACCTTTCCATGCTGAATGACAGTGTCTTTGG 972

QY 148 -----TyrMetValITrpGlyGlnGlyArgHisMetAspArgValLeuGlyArg 164

DB 973 ATCCCGCTTTATGCTCAAGGAGAGAGACGCTGCAATGGCTT---AATGCATTA 1029

QY 165 ThrTyArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrPhe 184

DB 1030 ATCCTTAAGAAACAGCTGAAGAGTGCAGACTGCTATCATCATGAGACT-TATTCAATC 1088

QY 185 ThrGluArgMetMetAlaTyrCysAspGlnIlePheGlnAspGluThrGlyLysAsn-Arg 204

DB 1089 TGTCAGAGGTTACTGGCT-----GACCAACAAGTGCCTCATCA 1127

QY 204 GATGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGlu 224

DB 1128 AAGACCC-----AGCACAGGCTCTCTCATGTAACACTGCTGCACACAGATT 1172

QY 224 uCyGlnGluIleValValTyrGlyMetValSerAspSerTyrCysArgGlyLysSerHis 244

DB 1173 TTGTGATGAATTCACCTGATATGGTCTGG-----CCCTTCCTTAAGATTTGAATGC 1226

QY 244 sProSerValProTyrHisTyrPheGlnLysGlyArgLeuAspGluCysGlnMetTyrLe 264

DB 1227 AAAAGCTGTGAATATCATTTACTAGAT-----GACTTGAATATGATACTT 1274

QY 264 uAlaHisGlnGlnAlaIleProArgSerAlaHisArgPheIleThrGluLysAlaValPhe 284

DB 1275 TTCCACGCCAAGCCCTCAGCAATGCAATTGAATTAAGAACTGTAATGTGTACACAA 1334

QY 284 rArgTrpAlaLysLys 289

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

QY 1335 CAGAGGACACTAA 1350

DB 1335 CAGAGGACACTAA 1350

```

FILING DATE: 17-July-1995
PRIOR APPLICATION DATA: PCT/EP94/04289
APPLICATION NUMBER: 22-December-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5959078man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET INFORMATION: BOER 1050.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2026 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 301..1377
US-08-972-498-1

Alignment Scores:
Pred. No.: 1,47e-12 Length: 2026
Score: 196.50 Matches: 71
Percent Similarity: 45.93% Conservative: 42
Best Local Similarity: 28.86% Mismatches: 99
Query Match: 12.08% Indels: 35
Gaps: 8
DB:

US-10-019-114A-7 (1-302) x US-08-972-498-1 (1-2026)
QY 59 TyrserservalProaspGlyLysProleuValArgInuProCysArgserCysAlaVal 78
Db 673 CACGACCTCTGCTGAGATTTCACCAATGAAAAACCGAGTTTAAGACCTGCTGCTT 732
QY 79 ValserSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluLeuAspSerAlaGlu 98
Db 733 GTTGAAACTCTGGCATCTACTAGACAGATGATGCGCAAGGAGATTGACAGTCACANT 792
QY 99 CysValPheArgMetAsnGlnAlaProThrValGlyPheGluAlaAspValGlyLysArg 118
Db 793 TTGTAATCAGGTGCAATCTAGCTCTGCTGTGAGATTGCTGCGGATGTGGGACTAAA 852
QY 119 SerThrLeuArgValValSerHisThrSerValProleuLeu-----LeuArgAsn 135
Db 853 TCAGATTATTATACATGACCAACCATCACTTGTGAGAGAGACATTGGAAGCTTCGGAT 912
QY 136 TyrSer-----HisTyrPheGlnLysAlaArgAspThrLeu----- 147
Db 913 GAGAGTGAACAGCAAAATTTGTGCATAGACTTTCATGCTGAATGACAGTCTCTTGG 972
QY 148 -----TyrMetValTyrGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArg 164
Db 973 ATCCCGCTTTATGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1029
QY 165 ThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPhe 184
Db 1030 ATCTTAAGAACACACTGAAGATGCGAACTGCTATCATCACTAGACT-TATTCAAGC 1088
QY 185 ThrGluArgMetMetAlaTyrCysAspGlnLeuPheGlnAspGlnThrGlyLysAsn-Arg 204
Db 1089 TGTCAAGAGTTTACTGCT-----GACCAACAAAGTGCCCATCA 1127
QY 204 GARGInserGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGluLe 224
Db 1128 AAGACCC-----AGCACAGGCTCTCTCATGTACACACAGGCGCACAGATT 1172
QY 224 uCysGluGluIleValAlaTyrGlyLysMetValSerAspSerTyrCysArgGluLysSerHis 244
Db 1173 TTGTATGAATTAATTCACCTGATAGGGTTCTGG-----CCCTTCCTCAAGGATTGAATGG 1226

```

```

QY 244 sProSerValProTyrHisTyrPheGluLysGlyArgLeuAspGlnCysGlnMetTyrLe 264
Db 1227 AAACCTGTGAATATCATCTACTAGAT-----GACTTGAATATAGTACTT 1274
QY 264 uAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluLysAlaValPheSe 284
Db 1275 TTCACAGCAAGACCTCACAGATGCCATTGAATTCAAACCCCTGAATGCTACACAA 1334
QY 284 rArgTrrAlaLysLys 289
Db 1335 CAGAGAGCACTAAAA 1350

RESULT 8
US-08-899-545-1
Sequence 1, Application US/08899545
Patent No. 6020201
GENERAL INFORMATION:
APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
TITLE OF INVENTION: Isolated Polysialyl Transferases,
TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
TITLE OF INVENTION: Production and Use
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,545
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/503,133
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6020201man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET INFORMATION: BOER 1050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
SEQUENCE CHARACTERISTICS:
LENGTH: 2026 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 301..1377
US-08-899-545-1

Alignment Scores:
Pred. No.: 1,47e-12 Length: 2026
Score: 196.50 Matches: 71
Percent Similarity: 45.93% Conservative: 42
Best Local Similarity: 28.86% Mismatches: 99
Query Match: 12.08% Indels: 35
Gaps: 8
DB:

US-10-019-114A-7 (1-302) x US-08-899-545-1 (1-2026)
QY 59 TyrserservalProaspGlyLysProleuValArgInuProCysArgserCysAlaVal 78

```

Db 673 CACAGCCTCTGCTGTAAGTTTCAACCAATGAAAAACCGCAGTTTAAGCCTGCTGTT 732  
 Qy 79 ValSerSerGlyGlnMetLeuGlySerGlyLeuGlyValAgluIleAspSerAlaGlu 98  
 Db 733 GTTGGAACCTGCTGCTACTAGACAGATGTGCGAAGAGATTGACAGTCAACAT 792  
 Qy 99 CyeValPheArgMetLeuGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArg 118  
 Db 793 TTGTATACGAGTGCATTAAGTCCCTGCTGTGATGATTTCTCGATGATGGGACATAAA 852  
 Qy 119 SerThrIleuArgValValSerHisThrSerValProleuLeu-----LeuArgAsn 135  
 Db 853 TCGATTTTATTAACATGAACCCATCGTGTGACAGAGACATTTGAGGCTTTCCGAAT 912  
 Qy 136 TyrSer-----TyrMetValTppGlyGlnGlyArgHisMetLeuAspArgValIleuGlyArg 147  
 Db 913 GAGAGTACAGACAAATTTGTGATAGACTTTCCATGCTGATGATGACAGTGCCTTTGG 972  
 Qy 148 -----TyrMetValTppGlyGlnGlyArgHisMetLeuAspArgValIleuGlyArg 164  
 Db 973 ATCCCGCTTTCATGTCGCAAGAGAGAGAGACGCGTGAATGGGTT--AATGCAATTA 1029  
 Qy 165 ThrTyrArgThrIleuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPhe 184  
 Db 1030 ATCCTTAAGAACAGCTGAAGTCCGCAACTGCTTATCATCATGAGACT-TATTCAATGC 1088  
 Qy 185 ThrGluArgMetLeuAlaTyrCysAspGlnIlePheGlnAspGluThrGlyValAsn-Ar 204  
 Db 1089 TGTCAAGGTACTGCTGCT-----GACCAACAAAGTGGCCATCA 1127  
 Qy 204 GARGGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGluLe 224  
 Db 1128 AAGACCC-----AGCACAGGCTCTCTCATGATACACACTGGCCACACAGATT 1172  
 Qy 224 uCysGlnGluIleValValTyrGlyMetValSerAspSerTyrCysArgGluYrsSerHis 244  
 Db 1173 TTGTGATGAATTCACCTGTATGGGTTCTGG-----CCCTTCCCTTAAGATTTGAATGG 1226  
 Qy 244 sProSerValProTyrHisTyrPheGlnGlyValArgLeuAspGluCysGlnMetTyrIle 264  
 Db 1227 AAAAGCTGAATATCATTAATACGAT-----GACTGTAATATATGATATCTT 1274  
 Qy 264 uAlaHisGlnGlnAlaProArgSerAlaHisArgPheIleThrGluValAlaValPheSe 284  
 Db 1275 TTCACAGCAGAACCCCTCACAGATGCGCATTAAGATCAAAACCTGATGTGTCACAA 1334  
 Qy 284 rArgTrrAlaValys 289  
 Db 1335 CAGAGGAGCCTAATAA 1350  
 RESULT 9  
 US-08-446-875-15  
 ; Sequence 15, Application US/08446875  
 ; Patent No. 5858751  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Paulson, James C.  
 ; APPLICANT: Wen, Xiaohong  
 ; APPLICANT: Livingston, Brian Duane  
 ; APPLICANT: Gillespie, William  
 ; APPLICANT: Keim, Sarge  
 ; APPLICANT: Burlingame, Alma L.  
 ; APPLICANT: Medzhiradzky, Katalin  
 ; TITLE OF INVENTION: Compositions and Methods for the  
 ; TITLE OF INVENTION: Identification and Synthesis of Stablytransferrases  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pome, Smith, Lande & Rose  
 ; STREET: 2029 Century Park East, Suite 3800  
 ; CITY: Los Angeles  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 90067  
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING-SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/446,875  
 FILING DATE: July 12, 1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/102,385  
 FILING DATE: August 4, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oldenkamp, David J.  
 REGISTRATION NUMBER: 29,421  
 REFERENCE/DOCKET NUMBER: 111-197  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (310) 788-5000  
 TELEFAX: (310) 277-1297  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1020 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHEICAL: NO  
 US-08-446-875-15  
 Alignment Scores:  
 Pred. No.: 8 36e-13 Length: 1020  
 Score: 194.50 Matches: 56  
 Percent Similarity: 44.79% Conservative: 30  
 Best Local Similarity: 29.17% Mismatches: 67  
 Query Match: 11.96% Indels: 39  
 DB: 2 Gaps: 6  
 US-10-019-114a-7 (1-302) x US-08-446-875-15 (1-1020)  
 Qy 73 CysArgSerCysAlaValValSerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAla 92  
 Db 415 TGCAGGCGCTGCGCCCGTGTGGGCAACTGCGGCACTGAGGAGCTTCTTATGGGCT 474  
 Qy 93 GluIleAspSerAlaGluCysValPheArgMetLeuGlnAlaProThrValGlyPheGlu 112  
 Db 475 GAGATGACAGTGCATTAAGTCTTCTCAGATGAAACAGCGCCACCGCAGGATTTGAA 534  
 Qy 113 AlaAspValGlyGlnArgSerThrLeuArgValValSerHisThrSerValProleuLeu 132  
 Db 535 GCTGATGTTGGACCAAGACCAACCATCTGTGATCCCTGAGAGCTTCCGGAGCTG 594  
 Qy 133 LeuArgAsnTyrSerHisTyrPheGlnGlyValArgAspThrLeuTyrMetValTrrGly 152  
 Db 595 CCACATATGTCAATGATTCCTGTCGCTTCAAG--ACATGCACTTGAATGG-- 648  
 Qy 153 GlnGlyArgHisMetLeuAspArgValIleuGlyGlyArgThrTyrArgThrLeuLeu----- 170  
 Db 649 -----GTGGTGGAGCGCCATCACACGCGGACCATTTCCACCATCAATCCCGGTT 699  
 Qy 171 -----GlnLeuThrArgMetTyrProGlyLeuGlnVal 181  
 Db 700 CCGCAAGATCAAGATGAACAGATTAAGATCTGATCTTACCAACCCAGCTTCATCAAG 759  
 Qy 182 TyrThrPheThrGluArgMetLeuAlaTyrCysAspGlnIlePheGlnAspGluThrGly 201  
 Db 760 TATGCTTTGACAAACGCGTGCAGAGGAC----- 789  
 Qy 202 LysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAla 221  
 Db 790 -----GGGCGATACCATTAACCGGATCTCTCCGTCATCTTCTCA 831  
 Qy 222 LeuGluLeuGlnGluIleValValTyrGlyMetValSerAspSerTyrCysArgGlu 241  
 Db 832 ATCATGTCTGCGATGAGTGAAGTGTACGGCTTGGGGGACAGAC-----AAGGG 885

Qy 242 LysSerHisProSerValProTyrHisTyrPheGlu 253  
 Db 886 AACCTGGCAC-----CACTACTGGGAG 906

## RESULT 10

US-08-446-875-1  
 ; Sequence 1, Application US/08446875  
 ; Patent No. 5858751

## GENERAL INFORMATION:

APPLICANT: Paulson, James C.  
 APPLICANT: Wen, Xiaohong  
 APPLICANT: Livingston, Brian Duane  
 APPLICANT: Gillespie, William  
 APPLICANT: Kelm, Sorge  
 APPLICANT: Burlingame, Alma L.  
 APPLICANT: Medzhiradzky, Katalin  
 TITLE OF INVENTION: Compositions and Methods for the  
 IDENTIFICATION AND SYNTHESIS OF SIALYLTRANSFERASES  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Poms, Smith, Lande & Rose  
 STREET: 2029 Century Park East, Suite 3800  
 CITY: Los Angeles  
 STATE: CA  
 COUNTRY: USA

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/446,875

FILING DATE: July 12, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/102,385

FILING DATE: August 4, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Oldenkamp, David J.

REGISTRATION NUMBER: 29,421

TELEPHONE: (310) 788-5000

TELEFAX: (310) 277-1297

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: porcine

TISSUE TYPE: liver, submaxillary glands

FEATURE:

NAME/KEY: CDS

LOCATION: 91..1119

OTHER INFORMATION: /product="porcine Gal Beta 1,3

OTHER INFORMATION: GAlNAc alpha 2,3 sialyltransferase"

US-08-446-875-1

Alignment Scores:

Pred. No.: 1.9e-12 Length: 1218  
 Score: 192.50 Matches: 64  
 Percent Similarity: 46.79% Conservative: 38  
 Best Local Similarity: 29.36% Mismatches: 82  
 Query Match: 11.84% Indels: 34  
 DB: 2 Gaps: 7

US-10-019-114a-7 (1-302) x US-08-446-875-1 (1-1218)

Qy 40 HisHisPheProThrGlySerArgPro-ThrValProGlyProLeuHisPheSerGlyTyr 59  
 Db 444 CACCAT-----CAGGAGCTGTTCAGGTGGTGGGACGTG----- 483  
 Qy 59 rSerSerValProAspGlyLysProLeuValArgGlu-----ProCysArgSerGly 76  
 Db 484 -----GACCCCTGCTGGAGAAAGAGCTGTGAGCTGGCGGCTG 524  
 Qy 76 sAlaValValSerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluLeaPse 96  
 Db 525 CGCGCTGTGGGCACTGGGCAAGCTGAAGAGCTCTTAAGGAGCTTGAAGGCAAGCTGG 584  
 Qy 96 rAlaGluCysValPheArgMetSerGlnAlaProThrValGlyPheGluAlaAspValG 116  
 Db 585 CCAAGACTTCGTGTGAGTGAAGCAAGGCTCCACAGAGAGGCTTGAAGGCAAGCTGG 644  
 Qy 116 yGlnArgSerThrLeuArgValAlaSerHisThrSerValProLeuLeuArgAsnTy 136  
 Db 645 GAGCAAGACCAACCACCATTTCTGTACCCGAGAGCTTCCGGAGCTGGCGAGAGGT 704  
 Qy 136 rSerHisTyrPheGlnHisAlaArgAspThrLeuTyrMetValTyrGlyGlnArgH 156  
 Db 705 CAGC-----ATGATCTGTGTCCTCCCTTCAAGACCAACGA 737  
 Qy 156 sMetAspArgValLeuGlyGlyArgThrTyrArgThrLeuLeuThrArgMetTy 176  
 Db 738 CCTGGAGTGGTATGATGACCGCCACACCGGACCATTCACACCTAGCTTCTGT 797  
 Qy 176 rProGlyLeuGlnValTyrThrPheThrGluArgMetMetAlaTyrCysAspGlnLeP 196  
 Db 798 CCCCCC---AAATCAAACTCAAAAAGAGAAAGATTCATTTATACCCGGCCTTCAT 854  
 Qy 196 eGlnAspGluThrGlyLysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTrrp 216  
 Db 855 CAAGTACGTCTTCAGACAGGTGGCTGACAGGCGCAGGCGCTGACCGCTCCACTGGATPCT 914  
 Qy 216 eThrMetLeuAlaLeuGlnLeuGlyGlnGluGluValValTyrGlyMetValSerAs 236  
 Db 915 CTCGGATCTTCTCCCTGCACATCTGTGACGAGGTGAGCTTGTATGGCTTTGGGGCGGA 974  
 Qy 236 pSerTyrCysArgGluLysSerHisProSerValProTyrHisTyrPheGlu 253  
 Db 975 CAGC-----AAGGGAACCTGGCAC-----CACTACTGGGAG 1005

RESULT 11  
 US-08-102-385G-1  
 ; Sequence 1, Application US/08102385G  
 ; Patent No. 5962294

GENERAL INFORMATION:  
 APPLICANT: Paulson, James C.  
 APPLICANT: Wen, Xiaohong  
 APPLICANT: Livingston, Brian Duane  
 APPLICANT: Gillespie, William  
 APPLICANT: Kelm, Sorge  
 APPLICANT: Burlingame, Alma L.  
 APPLICANT: Medzhiradzky, Katalin  
 TITLE OF INVENTION: Compositions and Methods for the  
 IDENTIFICATION AND SYNTHESIS OF SIALYLTRANSFERASES  
 NUMBER OF SEQUENCES: 36  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Oppenheimer Wolff & Donnelly LLP  
 STREET: 2029 Century Park East, 38th Floor  
 CITY: Los Angeles  
 STATE: CA  
 COUNTRY: USA

ZIP: 90067  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/102,385G  
 FILING DATE: 04-AUG-1993  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/925369  
 FILING DATE: 04-AUG-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oldenkamp, David J.  
 REGISTRATION NUMBER: 29,421  
 REFERENCE/DOCKET NUMBER: 97-062  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (310) 788-5000  
 TELEFAX: (310) 277-1297  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1218 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHEICAL: NO  
 ORGANISM: porcine  
 TISSUE TYPE: liver, submaxillary glands  
 FEATURES:  
 NAME/KEY: CDS  
 LOCATION: 91..1119 /product= "porcine Gal Beta 1,3  
 OTHER INFORMATION: Galnac alpha 2,3 sialyltransferase"  
 US-08-102-385G-1

Alignment Scores:  
 Pred. No.: 1.9e-12 Length: 1218  
 Score: 192.50 Matches: 64  
 Percent Similarity: 46.79% Conservative: 38  
 Best Local Similarity: 29.36% Mismatches: 82  
 Query Match: 11.84% Indels: 34  
 DB: Gaps: 7

US-10-019-114A-7 (1-302) x US-08-102-385G-1 (1-1218)  
 QY 40 HSHHPheProHrghlySerArgPro-ThrValProGlyProLeuHshPheSerGlyTy 59  
 DB 444 CACCAT-----CAGGAGCTGTTCCAGGTGCTCGGAGACGTG-----483  
 QY 59 rSerSerValProApSpGlyLysProLeuValArgGlu-----ProCysArgSerCy 76  
 DB 484 -----GACCCCTGCTGGAGAAAGAGCTGCTACGCTGCGCGCTG 524  
 QY 76 bAlaValValSerSerSerGlyLysMetLeuGlySerGlyValaGluIleAps 96  
 DB 525 CCGCGTCTGTGGCACTCGGCAACCTGAAGAGTCTTACTATGCGCTCAGATGACAG 584  
 QY 96 rAlaGluCysValPheArgMetAenGlnAlaProThrValGlyPheGlnAlaApValG 116  
 DB 585 CACGACTTGTCTGTGAGTGAACAGGCCCCCAGAGGCTTGAAGCCGACGCTCGG 644  
 QY 116 yGlnArgSerThrLeuArgValValSerHshThrSerValProLeuLeuArgAenTy 136  
 DB 645 GACCAAGACACACCATTTCTGTATCCCCGAGAGCTTCCGAGAGTGGCGCAGAGGT 704  
 QY 136 rSerHshTyPheGlnLysAlaArgAenThrLeuTyMetValITrpGlyGlnGlyArgH 156  
 DB 705 CAGC-----ATGATCTGTCTGCTCCCTTCAAGACCAACGA 737  
 QY 156 sMetAaPArgValLeuGlyLysArgThrTyArgThrLeuLeuGlnLeuHrArgMetTy 176  
 DB 738 CCGAGGTGGGTATATCAGCCGACACACGAGCATCTCCACACCTAAGTTCTGT 797  
 QY 176 rProGlyLeuGlnValTyThrPheThrGluArgMetMetAlaTyCysAspGlnIlePh 196  
 DB 798 CCGCGCC---AAATCAAAAGTCAAAAAGAGAAATGCTGATTTATCACCGGCGCTTCA 854

QY 196 eGlnAaPArgValLeuGlyLysArgThrTyArgThrLeuLeuGlnLeuHrArgMetTy 216  
 DB 855 CAGTACGTTCTGACAGGTGCTGACGAGCCACGAGGCTTCACTGACCTGCT 914  
 QY 216 eTrnMetIleLeuAlaLeuGlnLeuCysGlnGluIleValValTyArgIleValSerAs 236  
 DB 915 CTCGATATCTTCTCTGACATCTGTGACAGAGTGTGACTTGTATGCTTTGGCGGGA 974  
 QY 236 pSerTyPArgGluLysSerHshProSerValProTyHshTyPheGlu 253  
 DB 975 CAGC-----AAAGGAAGTGCAC-----CACTACTGAGAG 1005

RESULT 12  
 US-08-626-994A-4  
 Sequence 4, Application US/08626994A  
 Patent No. 5798244  
 GENERAL INFORMATION:  
 APPLICANT: Shuichi TSUJI et al.  
 TITLE OF INVENTION: Sia'2,3Gal'1,4GlcNAc '2,8-  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Greenblum & Bernstein, P.L.C.  
 STREET: 1941 Roland Clarke Place  
 CITY: Reston  
 STATE: Virginia  
 COUNTRY: U.S.A.  
 ZIP: 20191  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 1.44 diskette  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: Word Perfect 5.1+ (ASCII)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/626,994A  
 FILING DATE: April 3, 1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 77469/1995  
 FILING DATE: April 3, 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Arnold Turk  
 REGISTRATION NUMBER: 33,094  
 REFERENCE/DOCKET NUMBER: P14595  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 716-1191  
 TELEFAX: (703) 716-1180  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1048 bases  
 TYPE: nucleic acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: linear  
 US-08-626-994A-4  
 Alignment Scores:  
 Pred. No.: 2.55e-12 Length: 1048  
 Score: 190.50 Matches: 65  
 Percent Similarity: 44.29% Conservative: 28  
 Best Local Similarity: 30.95% Mismatches: 88  
 Query Match: 11.72% Indels: 29  
 DB: Gaps: 7  
 US-10-019-114A-7 (1-302) x US-08-626-994A-4 (1-1048)  
 QY 60 SerSerValProApSpGlyLysProLeuValArgGluProCysArgSerCysAlaVal 79  
 DB 326 TCCCTGCTCCACATATGTCGCCCATTAATGAATGAATGTTTGTGCTGTGCT 385  
 QY 80 SerSerSerGlyLysMetLeuGlySerGlyLysGlyValaGluIleApsSerAlaGluCys 99  
 DB 386 GAAAGAGTGAATCTTGAACAGGAGTCACTGTGACAGAAATTAATGAATCAATTTT 445



```

Oy      Val PheArgMetSerGlnIlePheProThylValGlyPheGluAlaLeuAspValGlyGlnAsxSer 119
Db      446 GTTCTCGAGCAATTTTGCCGCCGACAGAGGGCTTTCCCAAAGAATGTGGGAAGAAAACC 505
Oy      120 ThrLeuArgValValSerHisThrSerValProLeuLeuLeuArgSerHisIleTy 139
Db      506 AACCTGCACA-----ACCTTCAAATCCGACATCTTAAGAAATTATTACCAACAT 553
Oy      140 PheGlnIlyAlaLArgSerHisIleutyRmetValILTrGlyGlnIlyArgHisIleMetAspArg 159
Db      554 CTTTAAACCATTCAGACCGTAAACAATCTTCTCAGTTTAAAAAACCTTGATGGGCC 613
Oy      160 ValLeu-----GlyGlyArgThrTyArgThrLeuLeu 170
Db      614 ATACTTTTGATCCCTGCGATTTTTCTTCCACACTCTGCAAGCTGAACAGAACCTGATG 673
Oy      171 GlnLeuThrArgMetTyPrproGly--LeuGlnValTyThrPheThrGlnArgMetMet 189
Db      674 GATTTTTTTGTGAGCACAGAGGTCAAGTTAAAGGTC-----CAGTGG 715
Oy      190 AlaTyrcysArgGlnIlePheGln-----AspGluThrGlyLyAsnArgArgGlnSer 207
Db      716 GCTTAGGCGCTGGAATATCATGCAACATGTCAACAGGTACTGGAAAAAACAACACTGTCA 775
Oy      208 GlySerPheLeuSerThrGlyTyTPheThrMetIleLeuAlaLeuGlnLeuCyGluGlu 227
Db      776 CCCAAACGACTGAGCACAGGTATCTTAATGTATACTCTTGACATCTGCAATGTGTGAAG 835
Oy      228 IleValIlyrGlyMetValSer-----AspSerTyCysArgGlySerHis 244
Db      836 ATCCACTGTGACGGTTTCTGCGCCCTTTGGATTTGACCCCAACAACAGGAG----- 888
Oy      245 ProSerValProTyHisIlyrPheGlnIly 254
Db      887 ---GATCTGCCCTACCACTACTATGACAAA 913

RESULT 13
US-08-957-742-4
; Sequence 4, Application US/08957742
; Patent No. 6017743
; GENERAL INFORMATION:
; APPLICANT: Shunichi TSUJI et al.
; TITLE OF INVENTION: Sia2,3gala1,4GlcNAc '2,8-
; TITLE OF INVENTION: STALYUTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,742
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/626,994
; FILING DATE: April 3, 1996
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180

```

```

: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
:     LENGTH: 1048 bases
:     STRANDEDNESS: not relevant
:     TOPOLOGY: linear
: US-08-957-742-4
:
Alignment Scores:
Pred. No.:          2,55e-12          Length: 1048
Score:             190.50             Matches: 65
Percent Similarity: 44.29%            Conservative: 28
Best Local Similarity: 30.95%         Mismatches: 88
Query Match:       11.72%             Indels: 29
DB:                3                  Gaps: 7
:
US-10-019-114A-7 (1-302) x US-08-957-742-4 (1-1048)
QY      60  SerSerValProArgGlyLysProLeuValArgGluProCysArgSerCysAlaValVal 79
Db      326  TCCCGCTCCCGAATGTCGCCCCCTTATGATAAGCCGTATTAAGTTGTCTGTGTT 385
QY      80  SerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluLeuAspSerAlaGluCys 99
Db      386  GGAAACAGTGAATTTGTGACAGGGAGTCAAGTGTGACAAAGAAATGATTAATCAGATTTT 445
QY      100 ValPheArgMetArgGlnAlaProThrValGlyPheGlnAlaAspValGlyGlnArgSer 119
Db      446  GTTTCCTCATGCAATTTTGCCCGCCAGACAGGCTTTCCAAAGATGTTGGAAGAAACCC 505
QY      120 ThrLeuArgValAlaSerSerIleThrSerValProLeuLeuLeuArgAsnIleSerSerIleTyr 139
Db      506  AACCTCAAC-----ACCTTCAATCCGAGCATCTTAGAGAAATATTAACAACAT 553
QY      140 PheGlnLysAlaArgAspThrIleuTyrMetValTyrGlyGlnGlyArgHisMetAspArg 159
Db      554  CTTTAAACCATTCAGAGCCGTAAACAATCTTCTCAGTTTAAAAAAGCTTGATGGGCC 613
QY      160 ValLeu-----GlyGlyArgThrTyrArgThrLeuLeu 170
Db      614  ATACTTGGAGTCCCTGCATTTTCTTCCACACTTCTGCACACTGTAAACGAAGCGTAAAG 673
QY      171 GlnLeuThrArgMetTyrProGly---LeuGlnValTyrThrPheThrGluArgMetMet 189
Db      674  GATTTTGTGTTGACACACAGAGTCACTGAATTAAGTCTC-----CAGTTG 715
QY      190 AlaTyrCysAspGlnIlePheGln-----AspGluThrGlyLysAsnArgArgGlnSer 207
Db      716  GCTTGGCCCTGGAATATATCATGCAACATGTCAACAGTACTGGAAAAACAAACCTGTCA 775
QY      208 GlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGluLeuCysGluGlu 227
Db      776  CCCAAACGACTAGACACAGCATGATCCAAATGTAATACCTTGATTCGATCACTGCAATATGTGAAGAG 835
QY      228 IleValIleTyrGlyMetValSer-----AspSerTyrCysArgGluLysSerHis 244
Db      836  ATCCACTTGTACGGTTTCTGGCCCTTTGGATTGACCCCAACACCCAGGAG----- 886
QY      245 ProSerValProTyrHisTyrPheGluLys 254
Db      887  ---GATTCGCCCTTACCCTACTATGACAAA 913
:
RESULT 14
US-08-626-994A-2
: Sequence 2, Application US/08626994A
: Patent No. 5798244
:
GENERAL INFORMATION:
: APPLICANT: SHUICHI TSUJI et al.
: TITLE OF INVENTION: SIA 2,3GLA 1,4GLCNAC, 2,8-
: TITLE OF INVENTION: SIALYLTRANSFERASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenblum & Bernstein, P.L.C.

```





```
Db      570 GGAACAGTGGAACTTGTGACAGGAGTGCAGTGCACAGAAATAGATAATCAGATTTT 629
QY      100 ValPheArgMetLeuGlnAlaProThrValGlyPheGlnAlaAspValGlyGlnArgSer 119
Db      630 GTTCTCTGAGCAATTGTGCCGACAGAGGCTTTCCAAAGATGTGGAGAGAAACC 689
QY      120 ThrLeuArgValValSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyr 139
Db      690 AACCTCACA-----ACCTTCATCCGAGCATCTTAGAGAAATATTACAACAAT 737
QY      140 PheGlnLysAlaArgAspThrLeuTyrMetValTyrGlyGlnGlyArgHisMetAspArg 159
Db      738 CTTTACCATTCAGAACCGTAACAACCTCTCCTCAGTTTAAAAAAGCTTGATGGGGCC 797
QY      160 ValLeu-----GlyGlyArgThrTyrArgThrLeuLeu 170
Db      798 ATACTTTGGATCCCTGCATTTTCTTCCACACTTTCGCAACTGTACGAGAACGCTTAGTG 857
QY      171 GlnLeuThrArgMetTyrProGly---LeuGlnValTyrThrPheThrGluArgMetLeu 189
Db      858 GATTTTGTGTGAGCACAGAGTCAGTTAAAGTC-----CAGTTG 899
QY      190 AlaTyrCysAspGlnIlePheGln----AspGluThrGlyLysAsnArgArgGlnSer 207
Db      900 GCTTGCGCTGGAATATCATGCAACATGTCACACAGTACTGGAAAAACAAACACTGTCA 959
QY      208 GlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGluLeuCysGluGlu 227
Db      960 CCCAAGAGACTGACACAGGATATCTTAATGTAATCTTTCATCTGCATATGTGAAGAG 1019
QY      228 IleValValTyrGlyMetValSer-----AspSerTyrCysArgGluLysSerHis 244
Db      1020 ATCCACTGTAGCGTTCTGSCCCTTGATTTGACCCCAACACACAGGAG----- 1070
QY      245 ProSerValProTyrHisTyrPheGluLys 254
Db      1071 ---GATCTGCCCTTACCACTACTATGACAAA 1097
```

Search completed: December 21, 2004, 22:52:56  
Job time : 113 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 21, 2004, 22:49:40 ; Search time 598 Seconds

(without alignments)

2788.115 Million cell updates/sec

Title: US-10-019-114a-7

Perfect score: 1626

Sequence: 1 MKRGRVLVILICSVFSAV.....FSRMAKRPVFAHPSMRT 302

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4093002 seges, 2760418825 residues

Total number of hits satisfying chosen parameters: 8186004

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h  
-Q=/cgn2\_1/USPTO.epool/US10019114/runat\_21122004\_100142\_3479/app.query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=tblsune62  
-TRANS=human40.cdt -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10019114@cgn\_1\_1\_480@runat\_21122004\_100142\_3479  
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1626	100.0	1556	11 US-09-978-360A-2	Sequence 2, Appl1
2	1626	100.0	1556	10 US-10-315-664-25	Sequence 25, Appl
3	1622	99.8	1735	10 US-09-895-298-15	Sequence 15, Appl
4	1605	98.7	1151	16 US-10-262-511-183	Sequence 183, App
5	861	53.0	743	9 US-09-823-245A-30	Sequence 30, Appl
6	812	49.9	966	10 US-09-895-298-121	Sequence 121, App
7	738	45.4	2956	18 US-10-844-874-19	Sequence 19, Appl
8	721	44.3	966	16 US-10-188-186-21	Sequence 21, Appl
9	721	44.3	3283	14 US-10-198-846-11536	Sequence 11536, A
10	720	44.3	3207	15 US-10-104-047-380	Sequence 380, App
11	700.5	43.1	1005	16 US-10-188-186-23	Sequence 23, Appl
12	684	42.1	1106	16 US-10-188-186-19	Sequence 19, Appl
13	639.5	39.3	2048	14 US-10-097-340-188	Sequence 188, App
14	639.5	39.3	2090	15 US-10-094-749-802	Sequence 802, App
15	636.5	39.1	2075	15 US-10-288-252-31	Sequence 31, Appl
16	636.5	39.1	2075	17 US-10-398-038-31	Sequence 31, Appl
17	618	38.0	2017	16 US-10-108-260A-895	Sequence 895, App
18	544.5	33.5	2401	10 US-09-946-374-55	Sequence 55, Appl
19	544.5	33.5	2401	13 US-10-052-586-205	Sequence 205, App
20	544.5	33.5	2401	14 US-10-174-590-205	Sequence 205, App
21	544.5	33.5	2401	14 US-10-176-758-205	Sequence 205, App
22	544.5	33.5	2401	14 US-10-175-737-205	Sequence 205, App
23	544.5	33.5	2401	14 US-10-174-581-205	Sequence 205, App
24	544.5	33.5	2401	14 US-10-176-483-205	Sequence 205, App
25	544.5	33.5	2401	14 US-10-176-749-205	Sequence 205, App
26	544.5	33.5	2401	14 US-10-176-914-205	Sequence 205, App
27	544.5	33.5	2401	14 US-10-176-915-205	Sequence 205, App
28	544.5	33.5	2401	14 US-10-173-706-205	Sequence 205, App
29	544.5	33.5	2401	14 US-10-173-738-205	Sequence 205, App
30	544.5	33.5	2401	14 US-10-175-752-205	Sequence 205, App
31	544.5	33.5	2401	14 US-10-176-757-205	Sequence 205, App
32	544.5	33.5	2401	14 US-10-176-752-205	Sequence 205, App
33	544.5	33.5	2401	14 US-10-176-913-205	Sequence 205, App
34	544.5	33.5	2401	14 US-10-180-552-205	Sequence 205, App
35	544.5	33.5	2401	14 US-10-180-557-205	Sequence 205, App
36	544.5	33.5	2401	14 US-10-173-700-205	Sequence 205, App
37	544.5	33.5	2401	14 US-10-174-572-205	Sequence 205, App
38	544.5	33.5	2401	14 US-10-174-579-205	Sequence 205, App
39	544.5	33.5	2401	14 US-10-174-582-205	Sequence 205, App
40	544.5	33.5	2401	14 US-10-174-588-205	Sequence 205, App
41	544.5	33.5	2401	14 US-10-175-739-205	Sequence 205, App
42	544.5	33.5	2401	14 US-10-175-740-205	Sequence 205, App
43	544.5	33.5	2401	14 US-10-175-743-205	Sequence 205, App
44	544.5	33.5	2401	14 US-10-176-488-205	Sequence 205, App
45	544.5	33.5	2401	14 US-10-176-492-205	Sequence 205, App

#### ALIGNMENTS

RESULT 1  
US-09-978-360A-2  
; Sequence 2, Application US/09978360A  
; Publication No. US20040110939A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
; APPLICANT: Ducleit, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Uobert, Severin  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: 56 US4.CIP  
; CURRENT APPLICATION NUMBER: US/09/978,360A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: US 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: US 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: US 60/081,563

```

? PRIOR FILING DATE: 1998-04-13
? PRIOR APPLICATION NUMBER: US 60/096,116
? PRIOR FILING DATE: 1998-08-10
? PRIOR APPLICATION NUMBER: US 60/099,273
? PRIOR FILING DATE: -09-04
? PRIOR APPLICATION NUMBER: US 09/191,997
? PRIOR FILING DATE: 1998-11-13
? PRIOR APPLICATION NUMBER: US 09/215,435
? PRIOR FILING DATE: 1998-12-17
? PRIOR APPLICATION NUMBER: PCT/IB98/02122
? PRIOR FILING DATE: 1998-12-17
? PRIOR APPLICATION NUMBER: US 09/247,155
? PRIOR FILING DATE: 1999-02-09
? Remaining Prior Application data removed - See File Wrapper or PALM
? NUMBER OF SEQ ID NOS: 810
? SOFTWARE: Patent.pm
? SEQ ID NO 2
? LENGTH: 1556
? TYPE: DNA
? ORGANISM: Homo Sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 261..1166
? FEATURE:
? NAME/KEY: sig_peptide
? LOCATION: 261..314
? OTHER INFORMATION: Von Heijne matrix
? OTHER INFORMATION: score 8.80
? OTHER INFORMATION: seq RLVITLLCSVFS/AV
? FEATURE:
? NAME/KEY: polyA_site
? LOCATION: 1524..1556
? US-09-978-360A-2

```

**Alignment Scores:**

Pred. No.:	1,17e+196	1556
Score:	1626.00	302
Percent Similarity:	100.00%	
Best Local Similarity:	100.00%	
Query Match:	100.00%	
DB:	11	
	Gaps:	0

US-10-019-114A-7 (1-302) X US-09-978-360A-2 (1-1556)

Qy	1	MetLysAlaProGlyIyrgrueValleuIellIleuCysserValValPheSerAlaVal	20
Db	261	ATMAAGGCTCCGGGTGGGCTCGTGCATCAACCTGTCTCCGGTCTTCTTCGCCGTC	320
Qy	21	TyrIleIeuLeuCyCysTrpAlaGlyLeuProIeuCybleuAlaThrCysLeuAspHis	40
Db	321	TACATCTCTGTGGCTGGCGGGCCGGCCCTGCTCTGCTGGCCACTGCTTGACAC	380
Qy	41	HisPheProthrGlySerArgProthrValProGlyProIeuHisPheSerGlyTyrSer	60
Db	381	CACCTTCCCAAGGCTCCAGGCCACTGTGGCCGGGAAACCTGTCACTTACGTGGATATAC	440
Qy	61	SerValProaspGlyLysPserProIeuValArgGluProCysArgSerCysAlaValValSer	80
Db	441	AGGTGTCCAGATGGGAAGCCGTGTGTCCGAGCCCTGTCCAGCTGTGGCTGTCTCC	500
Qy	81	SerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleaspSerAlaGluCysVal	100
Db	501	AGCTCCGGCCAAATGCTGGGGCTCAGGCCCTGGGGTGTAGATGTGACAGTGGCAATGGGTG	560
Qy	101	PheArgMetAsnGlnAlaProThrValGlyPheGluAlaaspValGlyGlnArgSerThr	120
Db	561	TTCGGCATGAAACAGGCGGCCAACCGTGGGCTTGAAGGGAGATGTGGCCACAGCAGACC	620
Qy	121	LeuArgValValSerHisThrSerValProIeuLeuLeuArgAsnTyrSerHisTyrPhe	140
Db	621	CTGGGTGTGCTTCACACACAGCGTGGCGGTGTGTGGCACTATTTCACACTACTTC	680
Qy	141	GlnLysAlaArgAspThrIeuTyrMetValTyrGlyGlnGlyAlaGHisMetAspArgVal	160

Db	681	CAGAGGCCCGGAGACACGCTTACATGCTGTGGGGCCAGGGACAGCACATGACCGGGTG	740
Qy	161	LeuGIyGIyArGTThrTyTArGTThrLeuLeuGIuIeuThrArGTMeTyTrrProGIyLeuGIu	180
Db	741	CTGGCGGGCGGACACTACCGGACCGCTCTCGACGTCCACGAGATGACCCCGGCGCTGAG	800
Qy	181	ValTyTrThrPheThrTrGIuArGTMeMeAlaTyTyrCysaspGnIIepheGIuAspGIuThr	200
Db	801	GTGTACACCTTACAGGAGCGCATGTATGAGCCCTACCTGACCAAGATCTTCCAGAGACAGACG	860
Qy	201	GIyIyAsnArTArGArGIuInSerGIySerPheIeuSerThrGIyTrPheThrMetIleIeu	220
Db	861	GGCAAGAACCGGAGGCGAGTCGGAGCTCTCTTCTCAGCACCGCGTGGTTCACCATGATCCTC	920
Qy	221	AlaLeuGIuIeuCysGIuGIuIleValTyTrGIyMetValSerAspSerTyTyrCysArg	240
Db	921	GGCGTGGAGCGTGTGTAGAGAGATGTGTGTATGATGGATGGTCAACGACAGCTACTCGACGG	980
Qy	241	GIuIySerThIspSerSerValProTyTrHisTyPheGIuIyGIyAlaGluAspGIuIyCys	260
Db	981	GAGAGAGAGCCACCCCTCAGTGCCTTACACACTTTCAGAAAGGGCCGGCTAGATGAGTGT	1040
Qy	261	GIuMetTyTrLeuAlaHisGIuGIuAlaProArgSerAlaHisArgPheIleThrGIuIyAs	280
Db	1041	CAGATGTACCTGGGACACAGACAGGACGGCCCCGAAAGCGCCACCGCTTCATACTGAGAAAG	11000
Qy	281	AlaValPheSerArGTTPAlaTyIyAsArgProIleValPheAlaHisProSerTyTyrArg	300
Db	1101	GGCGTCTTCTCCCGCTGGGCCCAAGAAAGGCCCATCGTGTCCGCCCATCGCTCCTCGAAGG	11600

## RESULT 2

US-10-315-664-25  
Semience 25  
A

Publication No. US20030203377A1

GENERAL INFORMATION:

APPLICANT: Dumas Mine Edwards, C.D.  
APPLICANT: Bouqueleret, L.

APPLICANT: Jobert, S.

TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal  
 TITLE OR INVENTION: Pentides

FILE REFERENCE: GENSET.050CP3

CURRENT APPLICATION NUMBER: US/10/315,664

; CURRENT FILING DATE: 2002-12-09  
 ; PRIOR APPLICATION NUMBER: US/09/

PRIOR FILING DATE: 2000-06-21

PRIOR APPLICATION NUMBER: 60/113

PRIOR FILLING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/141

PRIOR FILING DATE: 1999-06-25

PRIOR APPLICATION NUMBER: 09/469

NUMBER OF SEO ID NOS: 123

SOFTWARE: Patent.pm

! SEQ ID NO 25

LENGTH: 1536  
TYPE: DNA

ORGANISM: Homo Sapiens

**FEATURE:** **WASH/VIEW** - CDS

NAME/KEY: CDS  
LOCATION: 261...1166

FEATURE:

NAME/KEY: sig\_peptide  
LOCATION: 261-314

OTHER INFORMATION: Von Heijne m

OTHER INFORMATION: score 8.80

OTHER INFORMATION: seq RLVLILC

```
NAME/KEY: polyA site
```

LOCATION: 1524..1556  
US-10-315-664-25

## Alignment Scores:

Pred. No.: 1.17e-196 Length: 1556  
Score: 1626.00 Matches: 302  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-10-019-114a-7 (1-302) x US-10-315-664-25 (1-1556)

```

Qy 1 MetLysAlaProGlyArgLeuValLeuIleIleLeuCySerValValPheSerAlaVal 20
Db ATAAAGGCTCCGGGCTCGGCTCGTATCATCTGTGCTCCGGGCTTCTTGGCGTC 320
Qy 21 TyrIleLeuLeuCySerValPheValLeuProLeuCySerValPheSerAlaVal 40
Db TACATCTCTGTGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 380
Qy 41 HisPheProThrGlySerValProThrValProGlyProLeuHisPheSerGlyTyrSer 60
Db CACTTCCCAACAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTATAGC 440
Qy 61 SerValProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValValSer 80
Db AGTGTCCAGATGAGAAAGCGGCTGTGCTCGGAGCGGCTGTGCTGTGCTGTGCTG 500
Qy 81 SerSerGlyGlnMetLeuGlySerGlyLeuGlyValGlnIleAspSerAlaGluCysVal 100
Db AGCTCCGGGCAATGCTGGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCG 560
Qy 101 PheArgMetAsnGlnAlaProThrValGlyPheGlnAlaAspValGlyGlnArgSerThr 120
Db TTCGGATGAACAGAGCGGCCACGTTGAGGCGGATGTGGGCGGAGCGGAGCACC 620
Qy 121 LeuArgValValSerIleThrSerValProLeuLeuLeuArgValTyrSerHisTyrPhe 140
Db CTGGGCTGTGCTCGACACAAAGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 680
Qy 141 GlnLysAlaArgAspThrLeuTyrMetValTrrGlyGlnIleArgHisMetAspArgVal 160
Db CAGAAAGCGGAGACAGCTCTACATGTGTGGGCGGAGGCGGAGCGGAGCGGAG 740
Qy 161 LeuGlyGlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGln 180
Db CTCGGGCGGCGACCTACCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 800
Qy 181 ValTyrThrPheThrGluArgMetMetAlaTyrCysAspGlnIlePheGlnAspGluThr 200
Db GTGTACACCTTCACGAGGCGGATGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCT 860
Qy 201 GlyLysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTrrPheThrMetIleLeu 220
Db GGGAAAGCGGAGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 920
Qy 221 AlaLeuGlnLeuCySerGlnIleValValTyrGlyMetValSerAspSerTyrCysArg 240
Db GCGCTGAGCTGTGTGAGGAGATCGTGTGTGTGAGGATGTGTGAGGATGTGTGAG 980
Qy 241 GlnLysSerHisProSerValProTyrHisTyrPheGlnLysGlyArgLeuAspGluCys 260
Db GAGAAAGCGGACCCCTCAGGCTTCACCTTGTGAGAAAGGCGGCTGTGAGTGTGT 1040
Qy 261 GlnMetTyrLeuAlaHisGlnIleAlaProArgSerAlaHisArgPheIleThrGluLys 280
Db CAGATTTACCTGACACAGGAGGCGGCGGCGGAGGCGGCGGCTTCATCAGTGAAG 1100
Qy 281 AlaValPheSerArgTrrAlaLysLysArgProIleValPheAlaHisProSerTrrArg 300
Db GCGGCTTCTCCCGGCTGGGCGGAGGAGGCGGATGTGTGTGCGGCTGTGAGAG 1160

```

Qy 301 ThrGlu 302  
Db 1161 ACTGAG 1166

## RESULT 3

```

US-09-895-298-15
; Sequence 15, Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: P2035P1
; CURRENT APPLICATION NUMBER: US/09/895,298
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591,16
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: PCT/US99/29590
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,006
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/112,809
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1735
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1002)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-895-298-15

```

## Alignment Scores:

Pred. No.: 4.45e-196 Length: 1735  
Score: 1622.00 Matches: 301  
Percent Similarity: 99.67% Conservative: 0  
Best Local Similarity: 99.67% Mismatches: 1  
Query Match: 99.75% Indels: 0  
DB: 10 Gaps: 0

US-10-019-114a-7 (1-302) x US-09-895-298-15 (1-1735)

```

Qy 1 MetLysAlaProGlyArgLeuValLeuIleIleLeuCySerValValPheSerAlaVal 20
Db ATAAAGGCTCCGGGCTCGGCTCGTATCATCTGTGCTCCGGGCTTCTTGGCGTC 352
Qy 21 TyrIleLeuLeuCySerValPheValLeuProLeuCySerValPheSerAlaVal 40
Db TACATCTCTGTGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 412
Qy 41 HisPheProThrGlySerValProThrValProGlyProLeuHisPheSerGlyTyrSer 60
Db CACTTCCCAACAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTCCAGGCTATAGC 472
Qy 61 SerValProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValValSer 80
Db AGTGTCCAGATGAGAAAGCGGCTGTGCTCGGAGCGGCTGTGCTGTGCTGTGCT 532
Qy 81 SerSerGlyGlnMetLeuGlySerGlyLeuGlyValGlnIleAspSerAlaGluCysVal 100
Db AGCTCCGGGCAATGCTGGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCG 592
Qy 101 PheArgMetAsnGlnAlaProThrValGlyPheGlnAlaAspValGlyGlnArgSerThr 120
Db TTCGGATGAACAGGCGGCCACGTTGAGGCGGATGTGGGCGGAGCGGAGCACCAC 652
Qy 121 LeuArgValValSerHisThrSerValProLeuLeuLeuArgValTyrSerHisTyrPhe 140
Db CTGGGCTGTGCTCGACACAAAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCT 712
Qy 141 GlnLysAlaArgAspThrLeuTyrMetValTrrGlyGlnIleArgHisMetAspArgVal 160

```

```

Db      713 CAGAAAGCCCGAAGACGCTCTACATGCTGTGGGGCCAGGCGACATGAGCCGGGTG 772
Qy      161 leuGlYgIYArghThrYrArghThrLeuGlYleuThzrArghMetYrProglYleuGln 180
Db      773 CTGGGGGGCGGACACCTACCCAGCGCTCTGACGTCCAGAGTGTATCCCGGCGCTGAG 832
Qy      181 ValYrThrPheThrGluArgMetMetAlaYrCyAspGlnIlePheGlnAspGluThr 200
Db      833 GTGTACACCTTCCAGGAGCCATGATGGCTTACCTGACCAAGATCTTCCAGGAGCGAGACG 892
Qy      201 GlyYAspArghArgGlnSerGlySerPheLeuSerThrGlyThrPheThrMetIleLeu 220
Db      893 GGCAGAGACCGGAGGCGAGTGGGGCTCTTCTCTAGACACCGGCTGTGTACCATGATCTTC 952
Qy      221 AlaLeuGluLeuGlyGlnGluIleValYrGlyMetValSerAspSerYrCyAspArg 240
Db      953 GCGCTGGAGGCTGTGTAGAGAGATCGTGTCTATGGGATGGTTCAGCGACATCTGTCAGG 1012
Qy      241 GluYsSerHisProSerValProYrHisYrPheGluYsglyYrGluAspGluCys 260
Db      1013 GAGAGAGCCACCCCTCAGTGGCTTACCTACTTGTAGAGAGGCGCGCTAGATGACTGT 1072
Qy      261 GluMetYrLeuAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluYs 280
Db      1073 CAGATGACCTGGACAGAGAGGCGCGCCGAGAGCCCGACCTTCACTCATCTGAGAGAG 1132
Qy      281 AlaValPheSerArgTrpAlaYsYsArgProIleValPheAlaHisProSerTrpArg 300
Db      1133 GCGGCTCTTCCCGCTGGGCGAAGAGAGGCCATCGTGTTCGCCATCTCCTGAGAG 1192
Qy      301 ThrGlu 302
Db      1193 ACTGAG 1198

```

# RESULT 4 US-10-262-511-183

```

; Sequence 183, Application US/10262511
; Publication No. US20040038222A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glendene
; APPLICANT: Miller, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingtang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patnirajan, Meera
; APPLICANT: Szytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malvankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Raetelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Beghe, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262, 511
; CURRENT FILING DATE: 2003-05-28

```

```

; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 183
; LENGTH: 1151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (421)..(1074)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: wherein n may be a, c, g or t
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: wherein n may be a, c, g or t
; US-10-262-511-183

```

## Alignment Scores:

```

Pred. No.: 3,58e-194 Length: 1151
Score: 1605.00 Matches: 298
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.71% Indels: 0
DB: 16 Gaps: 0

```

US-10-019-114A-7 (1-302) x US-10-262-511-183 (1-1151)

```

Qy      5 GlyArgLeuValLeuIleIleLeuCysSerValValPheSerAlaValTyrIleLeuLeu 24
Db      181 GGTGGGCTGTGTCTCATCATCTGTGCTCGTGTCTCTCTCTCCGCTTACATCTCTCTG 240
Qy      25 CysCysTrpAlaGlyLeuProLeuCysLeuAlaThrCysLeuAspHisIspheProThr 44
Db      241 TGCTGCTGGGCGCGGCTGCTCTGCTGGCCACGCTGAGCACCACTTCCACACA 300
Qy      45 GlySerArgProThrValProGlyProLeuHisPheSerGlyTyrSerSerValProAsp 64
Db      301 GGCCTCCAGGCCCATGTCGCGGAGCCCTGCATCTTCACTGATATAGCACTGTGCCAGAT 360
Qy      65 GlyYsPheProLeuValArgGluProCysArgSerCysAlaValYsSerSerGlyGln 84
Db      361 GGGAGAGCGGTGTGTGCGGAGCCCTGCGGAGCTGTGCTGTGTCTGAGCTCCGGCAA 420
Qy      85 MetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCysValPheArgMetAsn 104
Db      421 ATGCTGGGCTCAGAGCTGGTGTCTGAGATGAGACAGTCCAGAGTCCGCTTCCGCAATAC 480
Qy      105 GlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThrLeuArgValVal 124
Db      481 CAGGCGCCACCGTGGGCTTTGAGGCGGATGTGGGCGAGCGACACCTGCGTGTCTC 540

```

OY	125	SeRiStSRSeValProLeuLeuLeuArganTYrSerHisTYrPheGluValAlaArg	144
Db	541	TCACACAGACGGTCCCGCTGCTGTGGGCAACTTATTCACACTACTTCCAGAAAGCCCGCA	600
OY	145	AspThrLeuTYrMetValTYrPglYglnGlyArgHisSmetAspArgValLeuGlyGlyArg	164
Db	601	GACACGGCTCTCATGTGTGTGGGGCCAGGGGCAAGGCACATGGACCGGGTCTCGGCGGGCG	660
OY	165	ThrTYrArgTrpLeuLeuGlnLeuThrArgMetTYrProGlyLeuGlnValTYrThrPhe	184
Db	661	ACCTACCGGACCGCTCTGCACTGCACTCACAGATGTAACCCGGGCTGTGCAAGTGCACCTTC	720
OY	185	ThrGluArgMetMetAlaTYrCYsAspGlnIlePheGlnAspGluThrGlyLysAsnArg	204
Db	721	ACGGAGCCCAATGATGGCTCTACGCAACAGATCTTCCAGAGACAGAGACGGGCAAGAACGG	780
OY	205	ArgGlnSerGlySerPheLeuSerThrGlyTYrTrpPheThrMetIleLeuAlaLeuGlnLeu	224
Db	781	AGGCAAGTGGGGTCTCTTCTCTCAAGCACCGGCTGTTCACCAATGATCTTGGCTGTGAAGCTG	840
OY	225	CysGlnGlnIleValAlaValTYrGlyMetValSerAspSerTYrCYsArgGluLysSerHis	244
Db	841	TGTGAGGAGATCGTGCTCTATGAGGATGTCAAGCAGCACTACTGCAAGGAGAAAGGCCAC	900
OY	245	ProSerValProTYrHisTYrPheGluYsgIyAgleuAspGluCysGlnMetTYrLeu	264
Db	901	CCCTCAGGCTTATCACACTACTATTGAGAAAGGCGCGCTAGATGATGTCACATGTACACTG	960
OY	265	AlaHisGlnGlnAlaProArgSerAlaHisArgPheIleThrGlnGluAlaValPheSer	284
Db	961	GCACACGAGCAGGGGCGCCGAGAAAGGCCCAACGGCTTCATCATGAGAAAGCGGTCTTCC	1020
OY	285	ArgTrpAlaValGlyArgProIleValPheAlaHisProSerTrpArgThrGlu	302
Db	1021	CGCTGGGCAAGAAAGGCCCATCTGTTCGCCATTCGCTCTGGAGACATGAG	1074





```

      ; PRIOR FILING DATE: 2001-07-13
      ; PRIOR APPLICATION NUMBER: 60/373881
      ; PRIOR FILING DATE: 2002-04-19
      ; PRIOR APPLICATION NUMBER: 60/305673
      ; PRIOR FILING DATE: 2001-07-16
      ; Remaining Prior Application data removed - See File Wrapper or PALM.
      ; NUMBER OF SEQ ID NOS: 368
      ; SOFTWARE: Custom
      SEQ ID NO 21
      ; LENGTH: 966
      ; TYPE: DNA
      ; ORGANISM: Homo sapiens
      ; FEATURE:
      ; NAME/KEY: CDS
      ; LOCATION: (19)..(934)
      ; US-10-188-186-21

Alignment Scores:
Pred. No.: 1,74e-81 Length: 966
Score: 721.00 Matches: 139
Percent Similarity: 65.41% Conservative: 52
Best Local Similarity: 47.60% Mismatches: 73
Query Match: 44.34% Indels: 28
DB: 16 Gaps: 6

US-10-019-114A-7 (1-302) x US-10-188-186-21 (1-966)

QY 26 CysTTPAlaGlyLeuProLeuGlyLeuAla----- 35
Db 83 TGCTGG-----TGTGGCGCTTGTAATGAAGTAATTCCTCATGCTACTAA 130

QY 36 ThrCysLeuAap-----HisHapheProThrglySerAapProThr 49
Db 131 ACTGCTTTGGAAACCTGGTACAAAGTGAGTACCATTCCTCTACATCA----- 181

QY 50 ValPProGlyProLeuHisPheSer--GlyTyrSerSerValProAapGlyLeuProLeuV 69
Db 182 ---GGCGGCCCTTGGAACTCACTATGATGATCAATTAATGAAAGACACAAAGACCTTTGC 238

QY 69 AlaArgLuoProCysArgSerCysAlaValValSerSerSerGlyGlnMetLeuGlySerG 89
Db 239 AACTGGAC---TGTGACCTTTGTGCATAGTGTCAAACTCAGGTCAAGATGTTGGCCAGA 235

QY 89 IlyLeuGlyAlaGlnIleAapSerAlaGlnCysValPheArgHecAangInAlaProThrv 109
Db 296 AGGTGGGAATGAGATAGATCATGCTCTCTGATTTGGAGAAATGAACAAATGCCCAACA 355

QY 109 AlGlyPheGlnLalaPValGlyGlnArgSerThreukArgValValSerHisThSerV 129
Db 356 AAGGTTATGAAGAAATATGTCGCGCCGATGACCAATGATTTGAGATTGTGCCAATACAGG 415

QY 129 AlProLeuLeuLeuArgAntYrSerHisTyrPheGlnIlyAlaAapThrLeuTyrm 149
Db 416 TTCCTCTTTTGCTAATAAAACCTGATATATTTTTCAGAGAACGAATACATCAATTATAG 475

QY 149 EtValTTPGlyGlnIlyArgHisMetAapArgValLeuGlyGlyArgThrTyrArgThrl 169
Db 476 TTAATTGGGAGCCTTCCGCAATATGAGAAAGATGCAATGGCAATCGTTTACAAATG 535

QY 169 euleGlnLeuThrrgmectYrPProGlyLeuGlnValYrThrPheThrgInIlyArgMeln 189
Db 536 TGAATAAACAAGTTGGTATCTATCCGAATGCCCAATATATACGTACACAGAAACCGCA 595

QY 189 ecalATYrCysAapGlnIlePheGlnAapGlyThrglyLyAsnAArgAsgInSerGlyS 209
Db 596 TGAGTATCTGATGATGAGATTTTAAAGAAAGAACTGGAGAGACAGAGTCAGTCTGGCT 655

QY 209 erPheLeuSerThrglyTTPPheThrMetIleLeuAlaLeuGlnLeuGlyGlnIly 229
Db 656 CATATCTCAGACACAGGGGTGTTTACTTCTTCCGCGCATGACGCGCTGTATGGCATTC 715

QY 229 alValTTPGlyMetValSerAapSerTyrCysArgGlyLySerHisPProSerValProt 249

```

```

Db      716  AGCTGACGGAGTATGATTAATGACACTTACTGTAAAGACAGAAGGTTATGAAAAAGTCCCT  775
QY      249  YTHSYTRPheGluLysglYatgLeuAspGluCYGsimetYrLeuAlaHiGluGlnA  269
Db      776  ACCATTATTAATCAACAGAGAA---GATGAGCTGTATGAAATATTTCTTCATGAACATG  832
QY      269  IAProASerzLahIsaRphneiEThrGluLysAlaValPheSerArgTPRAlaLysL  289
Db      833  CCCCATATGCGGGGTCAATGAGTTTATCATCGAAAGAAAGATGTTTGCTTAATGGCCACAGA  892
QY      289  YSArgProILeValPheAlaHiSProSerTrp  299
Db      893  AGCAGAGATTAATATTTCACATCCAAACTCG  924

RESULT 9
US-10-198-846-11536
; Sequence 11536, Application US/10198846
; Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198, 846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11536
LENGTH: 3283
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 3283
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11536

Alignment Scores:
Pred. No.: 1,066-80 Length: 3283
Score: 721.00 Matches: 139
Percent Similarity: 65.41% Conservative: 52
Best Local Similarity: 47.60% Mismatches: 73
Query Match: 44.34% Indels: 28
Gaps: 6

US-10-019-114A-7 (1-302) x US-10-198-846-11536 (1-3283)
QY      26  CysTTPAlaGlyLeuProLeuCYsLeuAla-----  35
Db      184  TGCTGG-----TTGGCGCTTGTGTAATAAGATGAATTTCCATGCTACTAA  231
QY      36  ThrCYsLeuAsp-----HisH1sPheProThrGlySerArgProThr  49
Db      232  ACTGCTTTGGACAACTGTACAAAGTGAATACATTTCTCTTACACATACA-----  282
QY      50  ValProGlyProLeuHisPheSer--GlyTYrSerSerValProAspGlyLysProLeuV  69
Db      283  ---GGCGGCCCTTCGAACATCATGTGATACATAAATGTGAAGACACAAGAGCCTTTGG  339
QY      69  AlaArgGluProCYArGserCYsAlaValValSerSerSerGlyGlnMetLeuGlySerG  89
Db      340  AACTGAG---TGTCACCTTTGTGCGCATAGTCTCAACATCAGAGTCAAGATGGTTGGCCAGA  396
QY      89  LysLeuGlyValAGLileAspSerzLahIsaRphneiAGLysValPheArgMetArgnAlaProThrV  109
Db      397  AGGATGGGAATGAGATGATGATGATCTCTCTGATTTGGAGATGAAACAATGCCCCACCA  456

```

```

QY      109 aIGlyPheGluAlaSPValGlyGlnArgSerThrLeuArgValValSerHisThrSerV 129
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      457 AAGTATATGAAGAAGATGTGGCGCATGACATGATTCGATGGTTGTGTCATACAGCG 516
QY      129 aLProLeuLeuLeuAArgSerThrSerHisThrPheGlnLysAlaArgAspThrLeuTyrM 149
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      517 TTCCTCTTTTGGCTAAAGAACCCGTGATTTATTTTCAAGAGAGCAATACTACTATTATG 576
QY      149 eValTrpGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArgThrTyrArgThrL 169
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      577 TTTATTTGGGACCTTCCGCAATATGAGAAAGATGGCAATGGCATTCGTTTACAACATG 636
QY      169 euleuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrGluArgMet 189
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      637 TGAAGAAAGACAGTTGATCTATCCGAATGCCAAATATACGTGACACAGAGAAACGCA 696
QY      189 eValAlaTyrCysArgGlnIlePheGlnAspGluThrGlyLysAsnArgArgGlnSerGlyS 209
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      697 TGAAGTATCTGTGATGAGGTTTATTAAGAGAGAACTGGAGAGACAGAGTCCAGTCTGGCT 756
QY      209 ePheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGlyLeuGlyGlnGluIleV 229
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      757 CATATCTCACACACAGGGTGGTTTACCTTCCTTGGCCATGACGCTGTTATGCGCATTC 816
QY      229 aValTyrGlyMetValSerAspSerTyrCysArgGlyLysSerHisProSerValProT 249
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      817 ACGCTCAAGGATATGATTAATGACACCTACCTGCAGACAGAGAGGATATAGAAAAGTCCCT 876
QY      249 YrHisTyrPheGlnLysGlyArgLeuAspGluCysGlnMetTyrLeuAlaHisGlnGlnI 269
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      877 ACCATTATTTATGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
QY      269 lAProArgSerAlaHisArgPheIleThrGlyLysAlaValPheSerArgTrpAlaLysL 289
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      934 CCCCATATGGGGGTCATATGAGTTTATCTGTAAGAAAGAGTGTTCATAAATGGGCCAAGA 993
QY      289 YSAArgProIleValPheAlaHisProSerTrp 299
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      994 AGCAGAGATATATATTTACATCCAAACTGG 1025

RESULT 10
US-10-047-380
/ Sequence 380, Application US/10104047
/ Publication No. US20030236392A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: NO. US20030236392A1el full length cDNA
/ FILE REFERENCE: H1-A0105
/ CURRENT APPLICATION NUMBER: US/10/104,047
/ PRIOR FILING DATE: 2002-03-25
/ PRIOR APPLICATION NUMBER:
/ NUMBER OF SEQ ID NOS: 4096
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 380
/ LENGTH: 3207
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-047-380

Alignment Scores:
Pred. No.: 1,386-80 Length: 3207
Score: 720.00 Matches: 139
Percent Similarity: 65.41% Conservative: 52
Best Local Similarity: 47.60% Mismatches: 73
Query Match: 44.28% Indels: 28
DB: 15 Gaps: 6

US-10-019-114a-7 (1-302) x US-10-104-047-380 (1-3207)

QY      26 CysTrpAlaGlyLeuProLeuGlyLeuAla----- 35
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      213 TGCTGG-----TTGGCGCTTGTGAATGAAGTAATTTCCCATTTGCTACTATA 260

```

```

QY      36 ThrCysLeuAsp-----HisHisPheProThrGlySerArgProThr 49
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      261 ACTGCTTTGGACACACCTGGTACAAAGTGATACCAATTCCTCCACACATACA----- 311
QY      50 ValLProGlyProLeuHisPheSer--GlyTyrSerSerValProAspGlyLysProLeuV 69
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      312 ---GGCGGCCCTTGGAACTCACTATGATATCATATAAATGAAGACACAAAGCCTTGGC 368
QY      69 aLArgGluProCysArgSerCysAlaValValSerSerGlyGlnMetLeuGlySerG 89
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      369 AACTGAC---TGACACTTTGTGCCATATGTCACAACTCAGCTCAGATGTGGCCGCA 425
QY      89 lYLeuGlyValIleAspSerAlaGluCysValPheArgMetAsnGlnAlaProThrV 109
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      426 AGGTGGAAATGATGATATGATCATCTCTCGATTTGGAGATGAAACATGCCCCACCA 485
QY      109 aIGlyPheGluAlaSPValGlyGlnArgSerThrLeuArgValValSerHisThrSerV 129
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      486 AAGTATATGAAGAAGATGTGGCGCATGACATGATTCGATGGTTGTGTCATACAGCG 545
QY      129 aLProLeuLeuLeuAArgSerThrSerHisThrPheGlnLysAlaArgAspThrLeuTyrM 149
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      546 TTCCTCTTTTGGCTAAAGAACCCGTGATTTTCAAGAGAGCAATACTACTATTATG 605
QY      149 eValTrpGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArgThrTyrArgThrL 169
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      606 TTTATTTGGGACCTTCCGCAATATGAGAAAGATGGCAATGCCATCATTTACAACATGT 665
QY      169 euleuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrGluArgMet 189
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      666 TGAAGAAAGACAGTTGATCTATCCGAATGCCAAATATACCTGACACAGAGAAACGCA 725
QY      189 eValAlaTyrCysArgGlnIlePheGlnAspGluThrGlyLysAsnArgArgGlnSerGlyS 209
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      726 TGAAGTATCTGTGATGAGGTTTATTAAGAGAGAACTGGAGAGACAGAGTCCAGTCTGGCT 785
QY      209 ePheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGlyLeuGlyGlnGluIleV 229
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      786 CATATCTCACACACAGGGTGGTTTACCTTCCTTGGCCATGACGCTGTTATGCGCATTC 845
QY      229 aValTyrGlyMetValSerAspSerTyrCysArgGlyLysSerHisProSerValProT 249
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      846 ACGTCTCAAGGATATGATTAATGACACCTACTGCAGAGAGAGGATATAGAAAAGTCCCT 905
QY      249 YrHisTyrPheGlnLysGlyArgLeuAspGluCysGlnMetTyrLeuAlaHisGlnGlnI 269
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      906 ACCATTATTTATGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
QY      269 lAProArgSerAlaHisArgPheIleThrGlyLysAlaValPheSerArgTrpAlaLysL 289
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      963 CCCCATATGGGGGTCATATGAGTTTATCTGTAAGAAAGAGTGTTCATAAATGGGCCAAGA 1022
QY      289 YSAArgProIleValPheAlaHisProSerTrp 299
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1023 AGCAGAGATATATTTACATCCAAACTGG 1054

RESULT 11
US-10-186-186-23
/ Sequence 23, Application US/10186186
/ Publication No. US20040029789A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson et al.
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 21402-397C
/ CURRENT APPLICATION NUMBER: US/10/186,186
/ PRIOR FILING DATE: 2002-07-02
/ PRIOR APPLICATION NUMBER: 60/303046
/ PRIOR FILING DATE: 2001-07-05
/ PRIOR APPLICATION NUMBER: 60/360814
/ PRIOR FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: 60/303828
/ PRIOR FILING DATE: 2001-09-07

```



Best Local Similarity: 45.21% Mismatches: 80  
 Query Match: 42.07% Indels: 28  
 DB: 16 Gaps: 6

US-10-019-114a-7 (1-302) x US-10-188-186-19 (1-1106)

```

QY 26 CysTrpAlaGlyLeuProLeuCySLeuAla----- 35
DB 106 TCGTGG-----TTGTGCGCTTGTAAATGAAGATTTCCATGCTACTAA 153
QY 36 ThrCysLeuAsp-----HisHisPheProThrGlySerArgProThr 49
DB 154 ACTGCTTTGACAACTGCTACAAAGTGAATGATCTCTCCACATACA----- 204
QY 50 ValProGlyProLeuHisPheSer--GlyTyrSerSerValProAspGlyLysProLeu 69
DB 205 ---GGCGGCCCCCTTGCACACTCATGATGATCAATAAATGGAAGACACAGACCTTTGC 261
QY 69 AlaArgLupProCysArgSerCysAlaValAlaSerSerGlyGlnMetLeuGlySerG 89
DB 262 AACTGGAC---TGTGACCTTTGTGTCATAGTCAAACTCAGGTGATGTCGCGCAGA 318
QY 89 LysLeuGlyAlaGluIleAspSerAlaGluCysValPheArgMetAsnGlnAlaProThr 109
DB 319 AGGTGGGAAATGAGATGATCGATCTCCCTGCTTGGAGAAATGAACATGCCCCACCA 378
QY 109 AlGlyPheGlnAlaAspValGlyGlnArgSerThrLeuArgValAlaSerHisThrSery 129
DB 379 AAGTTTGAAGAAGATGTCGCGCGCATGACATGATTCGAGTTGTGTCATACCGCGC 438
QY 129 AlProLeuLeuLeuArgAsnTyrSerHisTyrPheGlnLysAlaArgAspThrLeuTyr 149
DB 439 TTCCCTCTTTGTGCTAAAAACCCCTGATTAATTTTCAAGGAAGCCAAATACTACTATTA 498
QY 149 eTValTPGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArgThrTyrArgThr 169
DB 499 TTATTTGGGACCTTTCCGCAATATGAGAAAGATGCGCATGCGATGTTTCAACATGT 558
QY 169 euleGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrGluArgMet 189
DB 559 TGAAGAAAGACAGTGTATCTATCGAATGCCAAATATAGTACACACAGAGAGCGCA 618
QY 189 eTAlaTyrCysAspGlnIlePheGlnAspGluThrGlyLysAsnArgArgGlnSerGly 209
DB 619 TGAGTTACTGTGAAGAGATTTTAAAGAAAGCTGGAGAACAGACAGTACAGATGACG 678
QY 209 ePheLeuSerThrGlyTyrPheThrMetCileLeuAlaLeuGlnLeuCysGlnGluIle 229
DB 679 TGTTCATTGATCAAGAGGTGTTTACATTCATTTGCGCATGAGCGCTGTATGCGCATTC 738
QY 229 AlValTyrGlyMetValSerAspSerTyrCysArgGluLysSerHisProSerValPro 249
DB 739 ACGTCAAGCGAGTGAATATGACACCTACTGCAAGACAGAGGTATGAAAGTCCCT 798
QY 249 YrHisTyrPheGlnLysGlyArgLeuAspGluCysGlnMetTyrLeuAlaHisGlnGln 269
DB 799 ACCATTATATGAAACAGAGAGA---GATGAGTGTGATGATATTTTCTTATGAACATG 855
QY 269 LArProArgSerAlaHisArgPheIleThrGlyLysAlaValPheSerArgTrpAlaLys 289
DB 856 CCCCATATGCGGCGCATAGTTTATCAGTGAAGAAAGAGTGTGTTCTAAATGCGCAGAGA 915
QY 289 YSArgProIleValPheAlaHisProSerTrp 299
DB 916 AGCACAGATATATATTACATCAACAACTGG 947

```

## RESULT 13

US-10-097-340-188  
 ; Sequence 188, Application US/10097340  
 ; Publication No. US20030087250A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: John MONAHAN  
 ; APPLICANT: Manjula GANNANAVARAPU

```

APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIRBY
APPLICANT: Gordon B. MILLIS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHWANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GIATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
PRIOR APPLICATION NUMBER: 2002-03-14
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: PaedSeq for Windows Version 4.0
SEQ ID NO 188
LENGTH: 2048
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-340-188

Alignment Scores:
Pred. No.: 1,326-70 Length: 2048
Score: 639.50 Matches: 125
Percent Similarity: 64.75% Conservative: 44
Best Local Similarity: 47.89% Mismatches: 89
Query Match: 39.33% Indels: 3
DB: 14 Gaps: 3

US-10-019-114a-7 (1-302) x US-10-097-340-188 (1-2048)
QY 44 ThrGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTyrSerSerVal 62
DB 365 ACCGACAGCGGCGCGGGGTCCTCCGCGGACCGCGGCATGAGCGATACCTCGGAGTG 424
QY 63 ProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValAlaSerSerSer 82
DB 425 GCGGACACAGACCCCTG---AAATGCACTGAGGAGGACTGTCCCTGTGGACCACTCA 481
QY 83 GlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCysValPheArg 102
DB 482 GGGCATCTGTGCACACTGCGCAAGGCTCCAGATGACACAGACAGTGTGTCAATCCGC 541
QY 103 MetAsnGlnAlaProThrValGlyPheGlnAlaAspValGlyGlnArgSerThrLeuArg 122
DB 542 ATGAATACGCCGCCACACGCGGTATGGCGTATGCGGCAATGCGACACAGCTGAGG 601
QY 123 ValValSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPheGlnLys 142
DB 602 GTCATCCGCAATTCACAGATTCAGAGATCTCCGCAACCGCATGACCTGCTCAACGTG 661
QY 143 AlArgAspThrLeuTyrMetValTPGlyGlnGlyArgHisMetAspArgValLeuGly 162

```

Db 662 AGCAGGGGACCGGTTTCATCTTCTGAGGGCCCGACGACTACATCGGCGGAGCGGACG 721  
Qy 163 G1YArgThrTyraGthrLeuLeuGlnLeuThraGmetTyProGlyLeuGlnValTyr 182  
Db 722 GGCAGGCTCAACAACACTGCATCTCTGAGCAGGTCGCCCCGCTGAAGGCTTCC 781  
Qy 183 ThraPheThrGluArgMetMetAlaTyCyaspGlnIlePheGlnaPrgIuThryGlyLys 202  
Db 782 ATGATTACTGCCCAAGAAGTGGCTGACGTTTGATGATGCTTTCAGACGAGAGACTGGCAAA 841  
Qy 203 AsnaArgGlnSerGlySerPheLeuSerThryGlyTrpPheThrMetIleLeuAlaLeu 222  
Db 842 GACAGGAGATATCAACACTGCTGCTCAGCAGCTGCTGTTTACAAAGACATTGCACTG 901  
Qy 223 GluLeuCyagGluGluIleValTyrGlyMetValSeraspSerTyCyargGlyLys 242  
Db 902 GAGCTGTGACAGGATCAATGTTTATGCGATGGTGGCCCCAGACTCTCGAGGATGCC 961  
Qy 243 SerHisProSerValProTyHisTyPheGluLysGlyArgLeuAspGluCyargGlnMet 262  
Db 962 AATCACCTTCAGACTTATCATTAATATGAACTTTTGACCTGATGAATGTAACATG 1021  
Qy 263 TyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGluValAla 281  
Db 1022 TACTCTCCCATAGACGAGGACGAGGACGATCATACCCCTTATCAAGAGAAACGA 1081  
Qy 282 ValPheSerArgTrpAlaLysLysArgProIleValPheAlaHisProSerTrpArgThr 301  
Db 1082 GTCTTTAAGAACTGGGACGAGCACTTCATTAATTCACCTTTTTCACAGACTGGAACCA 1141  
Qy 302 Glu 302  
Db 1142 GAA 1144

## RESULT 14

US-10-094-749-802  
; Sequence 802, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: MAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOKYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 802  
; LENGTH: 2090  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-094-749-802

## Alignment Scores:

Pred. No.: 1,37e-70 Length: 2090  
Score: 639.50 Matches: 125  
Percent Similarity: 64.75% Conservative: 44  
Best Local Similarity: 47.89% Mismatches: 89  
Query Match: 39.33% Indels: 3  
DB: 15 Gaps: 3

US-10-019-114A-7 (1-302) x US-10-094-749-802 (1-2090)

Qy 44 ThrGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTySerSerVal 62  
Db 425 ACCCAGAGGCGCCCGGGGCTCCCGCGGACCGCGGCATCGACGATACCTCGAGTGG 484  
Qy 63 ProAspGlyLysProLeuValArgGluProCyargSerCyAlaValValSerSerSer 82  
Db 485 GCGGACACCAAGCCCTG---AAATGCACTGACAGGAGACTGTGCGCTGTGACAGCTCA 541  
Qy 83 GlyGlnMetLeuGlnSerGlyLeuGlnAlaGlnIleAspSerAlaGluCyArgValPheArg 102  
Db 542 GGGCATCTGTGACAGACTCGGCAAGGCTCCCAATGACCAAGACAGAGTGTTCATCCG 601  
Qy 103 MetAsnGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThrLeuArg 122  
Db 602 ATGAATAGCGCCCCACACCGCGCTATGCGCGTGAAGTGGCAATGACACAGCTTGAAG 661  
Qy 123 ValValSerHisThrSerValProLeuLeuLeuArgHisTrpSerHisTyPheGlnLys 142  
Db 662 GTCATCGCGATTCACAGCATCCAGAGAGATCTCCGCAACCGCATGACCTGCTCAACGTG 721  
Qy 143 AlaArgAspThrLeuTyMetValTrpGlyGlnGlyArgHisMetLeuPheValLeuGly 162  
Db 722 AGCCAGGGCACCGGTTCATCTTCTGGGCCCGCAGCAGCTACATGCGCGGAGACGCAAG 781  
Qy 163 GlyArgThrTyraGthrLeuLeuGlnLeuThraGmetTyProGlyLeuGlnValTyr 182  
Db 782 GGCAGGCTCAACAACACTGCATCTCTGAGCAGGTCGCTGCCGCTGAAGGCTTCC 841  
Qy 183 ThraPheThrGluArgMetMetAlaTyCyaspGlnIlePheGlnaPrgIuThryGlyLys 202  
Db 842 ATGATTACTGCCCAAGAAGTGGCTGACGTTTGATGATGACTCTTTCAGCAGAGACTGGCAAA 901  
Qy 203 AsnaArgGlnSerGlySerPheLeuSerThryGlyTrpPheThrMetIleLeuAlaLeu 222  
Db 902 GACAGGAGATATCAACACTGCTGCTCAGCAGCTGCTGTTTACAAATGCAATGCACTG 961  
Qy 223 GluLeuCyagGluGluIleValTyrGlyMetValSeraspSerTyCyargGlyLys 242  
Db 962 GAGCTGTGACAGGATCAATGTTTATGCGATGGTGGCCCCAGACTCTCGAGGATGCC 1021  
Qy 243 SerHisProSerValProTyHisTyPheGluLysGlyArgLeuAspGluCyargGlnMet 262  
Db 1022 AATCACCTTCAGACTTATCATTAATATGAACTTTTGACCTGATGAATGTAACATG 1081  
Qy 263 TyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGluValAla 281  
Db 1082 TACTCTCCCATAGACGAGGACGAGGACGATCATACCCCTTATCAAGAGAAACGA 1141  
Qy 282 ValPheSerArgTrpAlaLysLysArgProIleValPheAlaHisProSerTrpArgThr 301  
Db 1142 GTCTTTAAGAACTGGGACGAGCACTTCATTAATTCACCTTTTTCACAGACTGGAACCA 1201  
Qy 302 Glu 302  
Db 1202 GAA 1204

## RESULT 15

US-10-288-252-31  
; Sequence 31, Application US/10288252  
; Publication No. US20030143686A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: LAL, Preeti G.

```

APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: BURFORD, Neil
APPLICANT: GANDHI, Ameena R.
APPLICANT: WARREN, Bridget A.
APPLICANT: YAO, Monique G.
APPLICANT: TRIBOULET, Catherine M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: LEE, Ernestine A.
APPLICANT: HAFALIA, April J.A.
APPLICANT: LU, Yan
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: SANJANMALA, Madhu S.
APPLICANT: DING, Li
FILE OF INVENTION: TRANSFERASES
FILE REFERENCE: PI-0241 USA
CURRENT APPLICATION NUMBER: US/10/288, 252
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: PCT US01/30424
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/252, 819
PRIOR FILING DATE: 2000-11-21
PRIOR APPLICATION NUMBER: US 60/249, 639
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/247, 931
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/246, 001
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/244, 025
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: US 60/238, 481
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/236, 523
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PERL Program
SEQ ID NO 31
LENGTH: 2075
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US2003014366A1 2778782CB1
US-10-288-252-31

```

```

Alignment Scores:
Pred. No.: 3,26e-70 Length: 2075
Score: 636.50 Matches: 124
Percent Similarity: 64.75% Conservative: 45
Best Local Similarity: 47.51% Mismatches: 89
Query Match: 39.15% Indels: 3
DB: 15 Gaps: 3

```

US-10-019-114A-7 (1-302) x US-10-288-252-31 (1-2075)

```

QY 44 ThrGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTyrSerSerVal 62
DB 411 ACCCAGCAGCGCCCGGGGTCCTCCCGGAGCCGCGGCACCTGACGAGTACTCGAGATG 470
QY 63 ProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValValSerSerSer 82
DB 471 GCGGACCAAGAGCCCTG---AAATGCACTGCAAGGACTGTGCTGTGTGACCAAGCTCA 527
QY 83 GlyMetLeuGlySerGlyLeuGlyValGluIleAspSerAlaGluCysValPheArg 102
DB 528 GGGCATCTGTCAGACAGTCGCGCAAGGCTCCAGATTGACCAAGAGAGTGTGTCATCCGC 587
QY 103 MetLeuGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThrLeuArg 122
DB 588 ATGATGACGCGCCCAAGCGGCTATGCGCTGACGTGGCCAAATCGCACCAAGCCTGAGG 647
QY 123 ValValSerHisThrSerValProLeuLeuArgAsnTyrSerHisTyrPheGlnLys 142

```

```

DB 648 GTCATGCGCATTCAGATCCAGATCCAGAGATCTCCGCAACCGCATATGACTGTCAACGCG 707
QY 143 AlaArgAspThrLeuTyrMetValTyrGlyGlnGlyArgHisMetAspArgValLeuGly 162
DB 708 AGCCAGGAGCCGCTGTCATCTTCTGGGCCCCAGCAGCTACATGCGGCGGAGCCGCAAG 767
QY 163 GlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyr 182
DB 768 GCGCAGTCTACACAACTGCATCTCCGACGAGGTCTGCGCCGCGCTGAAGGCTTC 827
QY 183 ThrPheThrGluArgMetMetAlaTyrCysAspGlnIlePheGlnAspGlnThrGlyLys 202
DB 828 ATGATTACTCGCCACAAGATGCTGCAAGTTGATGAGCTTCAAGCAGAGACTGGCAAA 887
QY 203 AsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeu 222
DB 888 GACAGAGAGATATCCAAACTTGCTCAGACACTGGCTGATTCATGACAAATTGCACTG 947
QY 223 GluLeuCysGluGluIleValValTyrGlyMetValSerAspSerTyrCysArgGluLys 242
DB 948 GAGCTCTGACAGATCAATGTTATGCGATGTCGCCCAAGACTTGTGACGAGATCCC 1007
QY 243 SerHisProSerValProTyrHisTyrPheGluLysGlyArgLeuAspGluCysGlnMet 262
DB 1008 AATCACCTTCAGTACTTATCATTTATATGACCTTTTGGACCTGATGAATGTACAAAG 1067
QY 263 TyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGluLysAla 281
DB 1068 TACCTTCCATGAGGAGGAGCAGGAGGCGAGCATCAACCGCTTTTTCACAGAAACGA 1127
QY 282 ValPheSerArgTyrAlaLysLysArgProIleValPheAlaHisProSerTyrParGthr 301
DB 1128 GTCTTTAAGACGTGGGACAGGACATTCATATTCATTTTTCACACCAAGCTGGAACCA 1187
QY 302 Glu 302
DB 1188 GAA 1190

```

Search completed: December 22, 2004, 00:16:15  
Job time : 610 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:16:28 ; Search time 39 Seconds  
(without alignments)  
513.539 Million cell updates/sec

Title: US-10-019-114A-7  
Perfect score: 1626  
Sequence: 1 MKAPGRVLVILLCVVFSAV.....FSRWAKKRPVFAHPSWRTE 302

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodaca/1/1aa/5A COMB .pep:\*  
2: /cgn2\_6/ptodaca/1/1aa/5B COMB .pep:\*  
3: /cgn2\_6/ptodaca/1/1aa/6A COMB .pep:\*  
4: /cgn2\_6/ptodaca/1/1aa/6B COMB .pep:\*  
5: /cgn2\_6/ptodaca/1/1aa/PCFUS COMB .pep:\*  
6: /cgn2\_6/ptodaca/1/1aa/backfile1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	* Query Match Length	DB ID	Description
1	1626	100.0	302 4 US-09-599-360B-75	Sequence 75, Appl
2	639.5	39.3	336 3 US-09-334-601-13	Sequence 13, Appl
3	494	30.4	95 4 US-09-621-976-3915	Sequence 3915, Ap
4	198.5	12.2	340 2 US-08-446-875-16	Sequence 16, Appl
5	196.5	12.1	359 1 US-08-503-133A-2	Sequence 2, Appl
6	196.5	12.1	359 2 US-08-576-775A-2	Sequence 2, Appl
7	196.5	12.1	359 2 US-08-972-498-2	Sequence 2, Appl
8	196.5	12.1	359 2 US-08-899-545-2	Sequence 2, Appl
9	191	11.7	343 2 US-08-446-875-2	Sequence 2, Appl
10	191	11.7	343 2 US-08-102-385G-2	Sequence 2, Appl
11	190.5	11.7	339 1 US-08-626-994A-3	Sequence 3, Appl
12	190.5	11.7	339 3 US-08-957-742-3	Sequence 3, Appl
13	190.5	11.7	364 1 US-08-626-994A-1	Sequence 1, Appl
14	190.5	11.7	364 2 US-08-957-742-1	Sequence 1, Appl
15	178	10.9	355 2 US-08-666-367B-6	Sequence 6, Appl
16	178	10.9	355 3 US-08-143-438-6	Sequence 6, Appl
17	178	10.9	356 2 US-08-666-367B-5	Sequence 5, Appl
18	178	10.9	356 3 US-08-143-438-5	Sequence 5, Appl
19	174.5	10.7	329 3 US-09-334-601-10	Sequence 10, Appl
20	173.5	10.7	329 3 US-09-334-601-21	Sequence 21, Appl
21	172	10.6	375 2 US-08-446-875-10	Sequence 10, Appl
22	172	10.6	375 2 US-08-102-385G-10	Sequence 10, Appl
23	171	10.5	376 2 US-08-666-367B-8	Sequence 8, Appl
24	171	10.5	376 3 US-08-143-438-8	Sequence 8, Appl
25	170	10.5	374 2 US-08-446-875-4	Sequence 4, Appl
26	170	10.5	374 2 US-08-102-385G-4	Sequence 4, Appl
27	168	10.3	90 3 US-09-334-601-20	Sequence 20, Appl

28	168	10.3	406 1 US-08-446-777-4	Sequence 4, Appl
29	168	10.3	767 1 US-08-446-777-6	Sequence 6, Appl
30	168	10.3	767 1 US-08-446-777-8	Sequence 8, Appl
31	166.5	10.2	90 3 US-09-334-601-22	Sequence 22, Appl
32	166.5	10.2	90 3 US-09-334-601-24	Sequence 24, Appl
33	165	10.1	600 4 US-10-140-002-462	Sequence 462, App
34	163.5	10.1	356 4 US-08-361-304A-2	Sequence 2, Appl
35	159	9.8	340 2 US-08-102-385G-18	Sequence 18, Appl
36	157.5	9.7	90 3 US-09-334-601-23	Sequence 23, Appl
37	148.5	9.1	90 3 US-09-334-601-25	Sequence 25, Appl
38	145.5	8.9	90 3 US-09-334-601-26	Sequence 26, Appl
39	143	8.8	329 1 US-07-991-587A-7	Sequence 7, Appl
40	143	8.8	329 1 US-08-309-985-7	Sequence 7, Appl
41	143	8.8	332 2 US-08-446-875-12	Sequence 12, Appl
42	143	8.8	332 2 US-08-102-385G-12	Sequence 12, Appl
43	143	8.8	384 3 US-09-334-601-8	Sequence 8, Appl
44	142.5	8.8	77 2 US-08-102-385G-29	Sequence 29, Appl
45	142	8.7	55 2 US-08-102-385G-15	Sequence 15, Appl

#### ALIGNMENTS

```

RESULT 1
US-09-599-360B-75
Sequence 75, Application US/09599360B
Patent No. 6548633
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Bougueleret, L.
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
FILE REFERENCE: GENSET.050C23
CURRENT APPLICATION NUMBER: US/09/599,360B
CURRENT FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: 60/113,686
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/141,032
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/469,099
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 123
SOFTWARE: Patent.pm
SEQ ID NO 75
LENGTH: 302
TYPE: PRT
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -18..-1
US-09-599-360B-75
Query Match 100.0%; Score 1626; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 3.4e-170;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MKAPGRVLVILLCVVFSAVNYILLCACAGLPCLATCLDHPPTGSRPTPGSLHRSYGS 60
1 LRVVSHSVPLLRNYSYFQKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMPGLQ 180
121 LRVVSHSVPLLRNYSYFQKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMPGLQ 180
121 LRVVSHSVPLLRNYSYFQKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMPGLQ 180
181 VYFTEPMAYCQOIFQDETGRKRRSGSFLSGFTMLALCEIIVVYGAVSDSYCR 240
181 VYFTEPMAYCQOIFQDETGRKRRSGSFLSGFTMLALCEIIVVYGAVSDSYCR 240
241 EKSHPSVPIYHFKKGLDEQWLAHQAPRSARHPTTEKAVSRNAKKRPVFAHPSWR 300

```



Db 241 EXHPSVPIYIFKRGRLDECOMYLAHQAPRSARFITEKAVSRMAKKPPIFAHPSWR 300  
QY 301 TE 302  
Db 301 TE 302

## RESULT 2

US-09-334-601-13  
; Sequence 13, Application US/09334601  
; Patent No. 6280989  
; GENERAL INFORMATION:  
; APPLICANT: Kapitonov, Dmitri  
; APPLICANT: Yu, Robert  
; TITLE OF INVENTION: NOVEL STALYITRANSFERASES  
; FILE REFERENCE: VCUIP-6  
; CURRENT APPLICATION NUMBER: US/09/334,601  
; CURRENT FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 94  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-334-601-13

Query Match 39.3%; Score 639.5; DB 3; Length 336;  
Best Local Similarity 47.9%; Pred. No. 1,1e-61;  
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPYTP-GPLHSGYSSVPDCKEIVREPCRSQAVVSSGGQMLGSGLAGAIDSAECVFR 102  
Db 64 TQCPGVAPGRPIDGLGVADHKPL-KMCRDCALVTSGLHSHSQSQQIDQTECVIR 122  
QY 103 MNQAPYGVFADYVGSRTLVNVSHTSVPLLRNYSHYFOKARDTLXVWGQGRHMDRVLG 162  
Db 123 MNDAPRTGVGRDVGNRSLRIANISSIORILRNHDLNANSQGVVFIWGPSSYMRDQK 182  
QY 163 GRTRYTLQITRMVPGIQTFTERRMAYCDQIFQDSTGKRRSGSFLSTGFTMLAL 222  
Db 183 GOVNNHLLLSQVLPRLKAFWITRHKMLQFDELFOKETGDKRKISNTWLSGTWFTMLIAL 242  
QY 223 ELCEIIVYVGVNSYGRKSHSPVPPHYHPEKRGRLDECOMYLAHQAPR-SARHFI TEKA 281  
Db 243 ELCDRIIVYGVNPPDFCRDPHPSVPYHYIEPFGPDECTMILSHERKSGSHRFFITEKR 302  
QY 282 VFSRWAKRPPIVFAHPSWRTE 302  
Db 303 VFKNMARTFNIHFQDPWKDE 323

## RESULT 3

US-09-621-976-3915  
; Sequence 3915, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent .dm  
; SEQ ID NO 3915  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -18...-1

NAME/KEY: UNSURE  
LOCATION: 59  
OTHER INFORMATION: Xaa = Ala,Thr  
NAME/KEY: UNSURE  
LOCATION: 56  
OTHER INFORMATION: Xaa = Arg,Ser  
NAME/KEY: UNSURE  
LOCATION: 54  
OTHER INFORMATION: Xaa = His,Pro  
US-09-621-976-3915

Query Match 30.4%; Score 494; DB 4; Length 95;  
Best Local Similarity 96.8%; Pred. No. 1,7e-46;  
Matches 92; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAPGRLVLIILCSVFSAVYIILCCWAGLPICLATCDHHFPTGSRPTVPGPLHRSQYS 60  
Db 1 MKAPGRLVLIILCSVFSAVYIILCCWAGLPICLATCDHHFPTGSRPTVPGPLHRSQYS 60  
QY 61 SVPDGKPLVREPCRSQAVVSSGGQMLGSGLAGAID 95  
Db 61 SVPDGKPLVREKXCXCVVSSGGQMLGSGLAGAID 95

## RESULT 4

US-08-446-875-16  
; Sequence 16, Application US/08446875  
; Patent No. 5838751  
; GENERAL INFORMATION:  
; APPLICANT: Paulson, James C.  
; APPLICANT: Wen, Xiaohong  
; APPLICANT: Livingston, Brian Duane  
; APPLICANT: Gillespie, William  
; APPLICANT: Keim, George  
; APPLICANT: Burlingame, Alma L.  
; TITLE OF INVENTION: Compositions and Methods for the  
; TITLE OF INVENTION: Identification and Synthesis of Stalyltrnsferases  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fome, Smith, Lande & Rose  
; STREET: 2029 Century Park East, Suite 3800  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 90067  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,875  
; FILING DATE: July 12, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/102,385  
; FILING DATE: August 4, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oldenkamp, David J.  
; REGISTRATION NUMBER: 29,421  
; REFERENCE/DOCKET NUMBER: 111-197  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (310) 788-5000  
; TELEFAX: (310) 277-1297  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 340 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-446-875-16



Query Match 12.2%; Score 198.5; DB 2; Length 340;  
Best Local Similarity 27.8%; Pred. No. 3.6e-13;  
Matches 55; Conservative 33; Mismatches 60; Indels 51; Gaps 6;

QY 73 CRSCAVSSSGQMLGSLGAIEIDSAECVFRMNOAPTVGFADVGQRSTLRVSHTSVPLL 132  
DB 139 CRCAVAVNGNLRSSSYGPEIDSHDFVLRNNKAPFGAFADVGTKTHLV----- 190  
QY 133 LRVYSHYFQKARPTLVVWVGQRHMD-----RVLGRTYRRL-----QUTRM 175  
DB 191 ---YPSFRILGDVSMILVFPKIDLEWVVSATTTGTSHTYIPVPAKIRVKODKILY 247  
QY 176 YPGIQTFTFRMAYCDQIFODETGKRRSGSFLSTGWTMLALCEIIVYGVMS 235  
DB 248 HPAPIKTVFDMVLOGH-----GRYPTGLISLVFSNHCDEVLYIGFGA 291

QY 236 DSYCREKSHPSVPHYFE 253  
DB 292 DS--KGNMH-----HYME 302

RESULT 5  
US-08-503-133A-2  
; Sequence 2, Application US/08503133A  
; Patent No. 5747326  
; GENERAL INFORMATION:  
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;  
; TITLE OF INVENTION: Isolated Polysialyl Transferases,  
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of  
; TITLE OF INVENTION: Production and Use  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felle & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/503.133A  
; FILING DATE: 17-JULY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/04289  
; FILING DATE: 22-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5747326man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: BOER 1050  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-503-133A-2

Query Match 12.1%; Score 196.5; DB 1; Length 359;  
Best Local Similarity 27.5%; Pred. No. 6.4e-13;  
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPGDKPLVREBRCSCAVVSSSGQMLGSLGAIEIDSAECVFRMNOAPTVGFADVGQR 118  
DB 125 HSLLPVSPKMNRRFKTCVAVVNGSGLIDSGCGKEIDSHNFVLRCLNAPVVEFADVGTK 184

QY 119 SLIRVSHTSVPLL-----LRVYS-----HYFQKARPTL-----YMWGQGRHMDRVLGGR 164  
DB 185 SDFITMNPVSVOQAFGCFRSESDRAKFTVRLSLMNLDSVLMIPAPVWAGGKHYEW-NAL 243  
QY 165 TYRLLQLTTRMYPQLQV-----YTFERMAYCDQIFODETGKRRSGSFLSTGWTM 218  
DB 244 ILNKLKRVTAIVSLRLIHAVRGVWLTNKPV-----KRP-----STGLIMY 285  
QY 219 ILALICEEIVYGVMSDSYCREKSHPSVPHYFEKGRLECOMYLAHEQAPSAHFIT 278  
DB 286 TLATRFCEDEIHLGFV--PPKDLNGKAVKHYD---DLKRYYSNAPSHPMPLEFKT 339

QY 279 EKAVFSRWAK 289  
DB 340 LNVLHNKALK 350

RESULT 6  
US-08-576-775A-2  
; Sequence 2, Application US/08576775A  
; Patent No. 5849904  
; GENERAL INFORMATION:  
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;  
; TITLE OF INVENTION: Isolated Polysialyl Transferases,  
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of  
; TITLE OF INVENTION: Production and Use  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felle & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/576.775A  
; FILING DATE: 21-December-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/503.133  
; FILING DATE: 17-JULY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/04289  
; FILING DATE: 22-December-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5849904man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: BOER 1050.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-576-775A-2

Query Match 12.1%; Score 196.5; DB 2; Length 359;  
Best Local Similarity 27.5%; Pred. No. 6.4e-13;  
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPGDKPLVREBRCSCAVVSSSGQMLGSLGAIEIDSAECVFRMNOAPTVGFADVGQR 118  
DB 125 HSLLPVSPKMNRRFKTCVAVVNGSGLIDSGCGKEIDSHNFVLRCLNAPVVEFADVGTK 184

QY 119 STRVSHSVPL--LNYS-----HYOKARDTL-----YMWGGRHMDRYLGR 164  
DB 185 SDFITNPSVYORAFGGFNNESDRAKEFVRLSMLNDSVLMIPAFMWKGGKKEVENV-NAL 243  
QY 165 TYRTLLQLTMYPGLOV-----YTFTERMAVCDQIFODETKNRRSGSFLSTGWTM 218  
DB 244 ILKNKLKVRATYPSRLIHAVRGWLTNKVPI-----KRP-----STGLMW 285  
QY 219 ILALICEEIVVGMVSDSYCREKSHPSVPHYFEKGRLECOMYLAHQAPSAHREFT 278  
DB 286 TLATRFCEDEHLVGFW--PPKDLNGKAVKYHYD---DLKRYFSNAPSHPHMLEFKT 339  
QY 279 EKAVFSRWAKK 289  
DB 340 LNVLNHRGALK 350

## RESULT 7

US-08-972-498-2  
; Sequence 2, Application US/08972498  
; Patent No. 5959078  
; GENERAL INFORMATION:  
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;  
; TITLE OF INVENTION: Isolated Polysialyl Transferases,  
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of  
; TITLE OF INVENTION: Production and Use  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/972,498  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/576,775  
; FILING DATE: 21-December-1995  
; APPLICATION NUMBER: 08/503,133  
; FILING DATE: 17-July-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/04289  
; FILING DATE: 22-December-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5959078man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: BOER 1050.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-972-498-2

Query Match 12.1%; Score 196.5; DB 2; Length 359;  
Best Local Similarity 27.5%; Pred. No. 6,4e-13;  
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;  
59 YSSVDPGKPLVREPCRSQAVSSSGOMLGSGLGAEIDSAECVFRMNOAPTVGFADVGQR 118

DB 125 HSLLPFVSPMKRRFRTCAVGNSSGILLDSGGCKEIDSHNFIKCNLAAPVVEFADVGTR 184  
QY 119 STRVSHSVPL--LNYS-----HYOKARDTL-----YMWGGRHMDRYLGR 164  
DB 185 SDFITNPSVYORAFGGFNNESDRAKEFVRLSMLNDSVLMIPAFMWKGGKKEVENV-NAL 243  
QY 165 TYRTLLQLTMYPGLOV-----YTFTERMAVCDQIFODETKNRRSGSFLSTGWTM 218  
DB 244 ILKNKLKVRATYPSRLIHAVRGWLTNKVPI-----KRP-----STGLMW 285  
QY 219 ILALICEEIVVGMVSDSYCREKSHPSVPHYFEKGRLECOMYLAHQAPSAHREFT 278  
DB 286 TLATRFCEDEHLVGFW--PPKDLNGKAVKYHYD---DLKRYFSNAPSHPHMLEFKT 339  
QY 279 EKAVFSRWAKK 289  
DB 340 LNVLNHRGALK 350

## RESULT 8

US-08-899-545-2  
; Sequence 2, Application US/08899545  
; Patent No. 6020201  
; GENERAL INFORMATION:  
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;  
; TITLE OF INVENTION: Isolated Polysialyl Transferases,  
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of  
; TITLE OF INVENTION: Production and Use  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/899,545  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/503,133  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 6020201man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: BOER 1050  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-899-545-2

Query Match 12.1%; Score 196.5; DB 3; Length 359;  
Best Local Similarity 27.5%; Pred. No. 6,4e-13;  
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;  
59 YSSVDPGKPLVREPCRSQAVSSSGOMLGSGLGAEIDSAECVFRMNOAPTVGFADVGQR 118  
DB 125 HSLLPFVSPMKRRFRTCAVGNSSGILLDSGGCKEIDSHNFIKCNLAAPVVEFADVGTR 184



Db 193 VPESEFELAEVS-----MILVFKTDLIEWISATTTGTSHTYVPA-KIRY 242

QY 184 FTRRMAVCOIQIDGDKRKROSGLSTGFTMTALCEIIVYGWSDSYCREKS 243

Db 243 KKEKILYHPAFIKYVDRLQGHGRYPSIGISVIFSALICBVDLYRGADS--KGMW 300

QY 244 HPSVPHYFE 253

Db 301 H-----HYWE 305

RESULT 11

US-08-626-994A-3

Sequence 3, Application US/08626994A

Patent No. 5798244

GENERAL INFORMATION:

APPLICANT: Shuichi TSUJI et al.

TITLE OF INVENTION: Sta' 2,3gaia 1,4GICNAC ' 2,8-

TITLE OF INVENTION: STALYTRANSPERASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Greenblum & Bernstein, P.L.C.

STREET: 1941 Roland Clarke Place

CITY: Reston

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 20191

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 diskette

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Word Perfect 5.1+ (ASCII)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/626,994A

FILING DATE: April 3, 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 77469/1995

FILING DATE: April 3, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Arnold Turk

REGISTRATION NUMBER: 33,094

REFERENCE/DOCKET NUMBER: P14595

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 716-1191

TELEFAX: (703) 716-1180

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

US-08-626-994A-3

Query Match 11.7%; Score 190.5; DB 1; Length 339;

Best Local Similarity 30.0%; Pred. No. 2.7e-12;

Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGLVREPCRCSCAVVSSGOMLGSLGAEIDSAECVFRMNAQPTVGFADVGORS 119

Db 105 SLDPVSPIMNKRYNVCVAVGNSGILTGSCGGEIDKSDVSRCNPAFTEAFHKDVGRT 164

QY 120 TLRVSHTSVPLLRNYSHYFOKARDTL-----YMWGQGRHMDRVLGRTYRT 168

Db 165 NLT-----TFNPISILEKRYNNLLITODRNNFPLSLKLDGAILM-----IPAFPH 211

QY 169 LLOLTRYMPLQVYTFTER-----MAYCDOIFO--DETKNRROSGSFLSTGFTMTL 220

Db 212 SATVTRTL-----VDFVEHGRQLKVQLAWPNTIMQHVNRKMKHLSPKRLSTGLIMYTL 267

QY 221 ALBLCEIIVYGVNS--DSYCREKSHPSVPHYFEK 254

Db 268 ASAICEIHLYGFWPFGDPNTR-----DLPYHYDK 300

RESULT 12

US-08-957-742-3

Sequence 3, Application US/08957742

Patent No. 6017743

GENERAL INFORMATION:

APPLICANT: Shuichi TSUJI et al.

TITLE OF INVENTION: Sta' 2,3gaia 1,4GICNAC ' 2,8-

TITLE OF INVENTION: STALYTRANSPERASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Greenblum & Bernstein, P.L.C.

STREET: 1941 Roland Clarke Place

CITY: Reston

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 20191

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 diskette

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Word Perfect 5.1+ (ASCII)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,742

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/626,994

FILING DATE: April 3, 1996

APPLICATION NUMBER: JP 77469/1995

FILING DATE: April 3, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Arnold Turk

REGISTRATION NUMBER: 33,094

REFERENCE/DOCKET NUMBER: P14595

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 716-1191

TELEFAX: (703) 716-1180

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

US-08-957-742-3

Query Match 11.7%; Score 190.5; DB 3; Length 339;

Best Local Similarity 30.0%; Pred. No. 2.7e-12;

Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGLVREPCRCSCAVVSSGOMLGSLGAEIDSAECVFRMNAQPTVGFADVGORS 119

Db 105 SLDPVSPIMNKRYNVCVAVGNSGILTGSCGGEIDKSDVSRCNPAFTEAFHKDVGRT 164

QY 120 TLRVSHTSVPLLRNYSHYFOKARDTL-----YMWGQGRHMDRVLGRTYRT 168

Db 165 NLT-----TFNPISILEKRYNNLLITODRNNFPLSLKLDGAILM-----IPAFPH 211

QY 169 LLOLTRYMPLQVYTFTER-----MAYCDOIFO--DETKNRROSGSFLSTGFTMTL 220

Db 212 SATVTRTL-----VDFVEHGRQLKVQLAWPNTIMQHVNRKMKHLSPKRLSTGLIMYTL 267

QY 221 ALBLCEIIVYGVNS--DSYCREKSHPSVPHYFEK 254

Db 268 ASAICEIHLYGFWPFGDPNTR-----DLPYHYDK 300

RESULT 13

US-08-626-994A-1

Sequence 1, Application US/08626994A

Patent No. 5798244

GENERAL INFORMATION:

```

; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ' 2,8-
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/626,994A
; FILING DATE: April 3, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
;
; US-08-626-994A-1
;
; Query Match 11.7%; Score 190.5; DB 1; Length 364;
; Best Local Similarity 30.0%; Pred. No. 3e-12;
; Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;
;
; QY 60 SSVPDGKPLVREPCRCGAVVSSGOMLGSGIAIDSAECVFRMNOAPTYGFADVGORS 119
; DB 130 SLDPDVPFIMKRYKYNVCAVVGNSGILTGSQCGEIDSDVSRNFAPTFAFHVDGRT 189
; QY 120 TLRVSTSVPLLRNYSHYFQKARDTL-----YVWGQGHMDRVLCGRYRT 168
; DB 190 NLT-----TFNPSTLEKYYNNLITQDRNPFSLKLDGAILW-----IPAFFHT 236
; QY 169 LLOUTRMYPGLQYTFTR-----MAYCDQIFQ--DEGKRNROSGSLSTGWFMTIL 220
; DB 237 SATVTRTL-----VDFVEHKGQLKVQLAMPGNIMQHVNRYKXNHLSPKRLSTGILMYTL 292
; QY 221 ALCEIIEIVVYGWVS---DSYCREKSHPSVPHYFEK 254
; DB 293 ASAICEIHLYGFMPFGPDNTR-----DLPYHYDK 325
;
; RESULT 14
; US-08-957-742-1
; Sequence 1, Application US/08957742
; Patent No. 6017743
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ' 2,8-
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia

```

```

; COUNTRY: U.S.A.
; ZIP: 20191
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,742
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/626,994
; FILING DATE: April 3, 1996
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
;
; US-08-957-742-1
;
; Query Match 11.7%; Score 190.5; DB 3; Length 364;
; Best Local Similarity 30.0%; Pred. No. 3e-12;
; Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;
;
; QY 60 SSVPDGKPLVREPCRCGAVVSSGOMLGSGIAIDSAECVFRMNOAPTYGFADVGORS 119
; DB 130 SLDPDVPFIMKRYKYNVCAVVGNSGILTGSQCGEIDSDVSRNFAPTFAFHVDGRT 189
; QY 120 TLRVSTSVPLLRNYSHYFQKARDTL-----YVWGQGHMDRVLCGRYRT 168
; DB 190 NLT-----TFNPSTLEKYYNNLITQDRNPFSLKLDGAILW-----IPAFFHT 236
; QY 169 LLOUTRMYPGLQYTFTR-----MAYCDQIFQ--DEGKRNROSGSLSTGWFMTIL 220
; DB 237 SATVTRTL-----VDFVEHKGQLKVQLAMPGNIMQHVNRYKXNHLSPKRLSTGILMYTL 292
; QY 221 ALCEIIEIVVYGWVS---DSYCREKSHPSVPHYFEK 254
; DB 293 ASAICEIHLYGFMPFGPDNTR-----DLPYHYDK 325
;
; RESULT 15
; US-08-666-367B-6
; Sequence 6, Application US/08666367B
; Patent No. 5854042
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: NOVEL SUGAR-CHAIN SYNTHETASE AND PROCESS FOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1

```

